

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: December 19, 2003, 14:26:39 / Search time 32 Seconds
(without alignments)
1960.965 Million cell updates/sec

Title: US-09-995-938a-7
Perfect score: 1765
Sequence: 1 MTSFGATSTSAIAAAAAAAAAA.....IHVGMEDELTGNGKARG 336

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 696363 seqs, 186758610 residues
Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Applications other
sees

Result No.	Score	Query Length	DB ID	Description
1	1765	100.0	12	US-09-995-938a-7
2	1758	99.6	12	US-09-995-938a-6
3	1526.5	86.5	12	US-09-995-938a-10
4	1519.5	86.1	12	US-09-995-938a-8
5	137.5	7.8	12	US-09-995-938a-20
6	131	7.4	12	US-09-995-938a-21
7	127.5	7.2	12	US-10-043-487-311
8	127.5	7.2	12	US-10-105-004-2
9	127.5	7.2	12	US-10-105-004-121
10	127.5	7.2	12	US-10-105-004-125
11	126.5	7.2	12	US-09-858-754-4
12	126.5	7.2	12	US-10-106-698-5571
13	126.5	7.2	12	US-10-105-004-5
14	124.5	7.1	12	US-10-032-585-7168
15	124	7.0	12	US-10-105-004-131
			12	US-10-294-804-2

16	124	7.0	12	US-09-801-368-108	Sequence 108, App
17	122	6.9	12	US-09-917-384-1	Sequence 1, Appl
18	122	6.9	12	US-09-917-383-1	Sequence 1, Appl
19	120.5	6.8	12	US-09-858-754-3	Sequence 3, Appl
20	120.5	6.8	12	US-10-000-864-8	Sequence 8, Appl
21	120	6.8	12	US-09-789-386-2	Sequence 2, Appl
22	120	6.8	12	US-09-758-140-6	Sequence 6, Appl
23	120	6.8	12	US-09-893-348-23	Sequence 23, Appl
24	120	6.8	12	US-09-972-599A-6	Sequence 6, Appl
25	120	6.8	12	US-10-060-036-71	Sequence 71, Appl
26	119.5	6.8	12	US-10-094-466-14	Sequence 14, Appl
27	116.5	6.6	12	US-10-155-400-1	Sequence 1, Appl
28	115.5	6.5	12	US-09-826-752-6	Sequence 6, Appl
29	114.5	6.5	12	US-10-029-386-34203	Sequence 34203, A
30	114.5	6.5	12	US-09-945-917-57	Sequence 57, Appl
31	114.5	6.5	12	US-09-845-917A-57	Sequence 57, Appl
32	114.5	6.5	12	US-09-945-917-58	Sequence 58, Appl
33	114.5	6.5	12	US-09-845-917A-58	Sequence 58, Appl
34	114.5	6.5	12	US-09-945-917-3	Sequence 3, Appl
35	114.5	6.5	12	US-09-845-917A-3	Sequence 3, Appl
36	114.5	6.5	12	US-09-945-917-4	Sequence 4, Appl
37	114.5	6.5	12	US-09-845-917A-4	Sequence 4, Appl
38	113.5	6.4	12	US-10-029-386-33139	Sequence 33139, A
39	113.5	6.4	12	US-10-171-311-83	Sequence 83, Appl
40	112	6.3	12	US-10-163-774-3	Sequence 74, Appl
41	111.5	6.3	12	US-10-338-777-40	Sequence 40, Appl
42	111	6.3	12	US-10-094-466-16	Sequence 16, Appl
43	110.5	6.3	12	US-09-780-053-2	Sequence 2, Appl
44	110.5	6.3	12	US-10-409-511-2	Sequence 2, Appl
45	110.5	6.3	12	US-10-145-396-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1				
US-09-995-938a-7				
Sequence 7, Application US/0995938A				
Publication No. US20030150026A1				
GENERAL INFORMATION:				
APPLICANT: JOANNE CHORY				
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID				
FILE REFERENCE: SALKINS, 046A				
CURRENT FILING DATE: 2001-11-27				
NUMBER OF SEQ ID NOS: 14				
SOFTWARE: FASTSEQ for Windows Version 4.0				
SEQ ID NO 7				
LENGTH: 336				
TYPE: PRT				
ORGANISM: ARABIDOPSIS THALIANA				
US-09-995-938a-7				
Query Match	100.0%	Score 1765	DB 12	Length 336
Best Local Similarity	100.0%	Pred. No. 3.4e-134		
Matches 336	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MTSDGATSTSAIAAAAAAAAAARRKPSWRERENRRRAVAAKITGTAAGDVNLP	60	Sequence 1, Appl
DB	1	MTSDGATSTSAIAAAAAAAAAARRKPSWRERENRRRAVAAKITGTAAGDVNLP	60	Sequence 1, Appl
QY	61	KHCNNNEVLALCVAAWVEEDCTTRKCKPLPGELAGTSSKVTPTSSONQSPFLSSAF	120	Sequence 6, Appl
DB	61	KHCNNNEVLALCVAAWVEEDCTTRKCKPLPGELAGTSSKVTPTSSONQSPFLSSAF	120	Sequence 6, Appl
QY	121	OSPSPYQVSPSSSPSPSGEPNNMSSPTFFELRNNGIPSSLPRLINSNCVPTBPV	180	Sequence 11, Appl
DB	121	OSPSPYQVSPSSSPSPSGEPNNMSSPTFFELRNNGIPSSLPRLINSNCVPTBPV	180	Sequence 11, Appl
QY	181	SPTSPKPKPLPWMSIAKOSMAIPNYFVAVASPASPTRHGHFTLATIPEC	240	Sequence 181, Appl
DB	181	SPTSPKPKPLPWMSIAKOSMAIPNYFVAVASPASPTRHGHFTLATIPEC	240	Sequence 181, Appl

Db 181 SPSPTSKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIPEC 240
; SEQ ID NO 10
; LENGTH: 335
; TYPE: PRT
; ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-10

Qy 241 DESDSSSTVDSGHWISFOKFAOQOQPSASWPTSPFNLVKPAPOQMSPTAAFOEIQSS 300
Db 241 DESDSSSTVDSGHWISFOKFAOQOQPSASWPTSPFNLVKPAPOQMSPTAAFOEIQSS 300

Qy 301 EFKFENSQVKKPWEGERIHVDGMEDELTLTGNGKARG 336
Db 301 EFKFENSQVKKPWEGERIHVDGMEDELTLTGNGKARG 336

RESULT 2
US-09-995-938A-6
; Sequence 6, Application US/09995938A
; Publication No. US20030150026A1
; GENERAL INFORMATION:
; APPLICANT: JOANNE CHORY
; APPLICANT: ZHIYONG WANG
; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
; TITLE OF INVENTION: HORMONE ACTION IN PLANTS
; FILE REFERENCE: SALKINS. 046A
; CURRENT APPLICATION NUMBER: US/09/995.938A
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 336
; TYPE: PRT
; ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-6

Query Match 99.6%; Score 1758; DB 12; Length 336;
Best Local Similarity 99.7%; Pred. No. 1.2e-133;
Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MTSFGATSTSAIAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGIAOGDYNLP 60
Db 1 MTSFGATSTSAIAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGIAOGDYNLP 60

Qy 61 KHCDDNEVLKALCYEAGWVVEEDGTTTKGCKPLPGELAGTSSRVTPYSSQNSPLSSAF 120
Db 61 KHCDDNEVLKALCYEAGWVVEEDGTTTKGCKPLPGELAGTSSRVTPYSSQNSPLSSAF 120

Qy 121 QSPISYQVSPSSSSSPSPSRGEPNNNSSTFFPFLRNGIIPSSLSLRISNCPVTPPV 180
Db 121 QSPISYQVSPSSSSSPSPSRGEPNNNSSTFFPFLRNGIIPSSLSLRISNCPVTPPV 180

Qy 181 SSPTSCKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIPEC 240
Db 181 SSPTSCKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIPEC 240

Qy 241 DESDSSSTVDSGHWISFOKFAOQOQPSASWPTSPFNLVKPAPOQMSPTAAFOEIQSS 300
Db 241 DESDSSSTVDSGHWISFOKFAOQOQPSASWPTSPFNLVKPAPOQMSPTAAFOEIQSS 300

Qy 301 EFKFENSQVKKPWEGERIHVDGMEDELTLTGNGKARG 336
Db 301 EFKFENSQVKKPWEGERIHVDGMEDELTLTGNGKARG 336

RESULT 3
US-09-995-938A-10
; Sequence 10, Application US/09995938A
; Publication No. US20030150026A1
; GENERAL INFORMATION:
; APPLICANT: JOANNE CHORY
; APPLICANT: ZHIYONG WANG
; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
; TITLE OF INVENTION: HORMONE ACTION IN PLANTS
; FILE REFERENCE: SALKINS. 046A
; CURRENT APPLICATION NUMBER: US/09/995.938A
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 335
; TYPE: PRT
; ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-10

Query Match 86.5%; Score 1526.5; DB 12; Length 335;
Best Local Similarity 88.7%; Pred. No. 5.6e-115;
Matches 298; Conservative 10; Mismatches 23; Indels 5; Gaps 4;

Qy 1 MTSFGATSTSAIAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGIAOGDYNLP 60
Db 1 MTSFGATSTSAIAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGIAOGDYNLP 59

Qy 61 KHCDDNEVLKALCYEAGWVVEEDGTTTKGCKPLPGELAGTSSRVTPYSSQNSPLSSAF 120
Db 61 KHCDDNEVLKALCYEAGWVVEEDGTTTKGCKPLPGELAGTSSRVTPYSSQNSPLSSAF 119

Qy 121 QSPISYQVSPSSSSSPSPSRGEPNNNSSTFFPFLRNGIIPSSLSLRISNCPVTPPV 179
Db 121 QSPISYQVSPSSSSSPSPSRGEPNNNSSTFFPFLRNGIIPSSLSLRISNCPVTPPV 177

Qy 180 VSSPTSCKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIPEC 238
Db 178 VSSPTSCKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIPEC 237

Qy 239 ECDSDSSSTVDSGHWISFOKFAOQOQPSASWPTSPFNLVKPAPOQMSPTAAFOEIQSS 298
Db 238 ECDSDSSSTVDSGHWISFOKFAOQOQPSASWPTSPFNLVKPAPOQMSPTAAFOEIQSS 297

Qy 299 SSEFKFENSQVKKPWEGERIHVDGMEDELTLTGNGKARG 334
Db 298 SSEFKFENSQVKKPWEGERIHVDGMEDELTLTGNGKARG 333

RESULT 4
US-09-995-938A-8
; Sequence 8, Application US/09995938A
; Publication No. US20030150026A1
; GENERAL INFORMATION:
; APPLICANT: JOANNE CHORY
; APPLICANT: ZHIYONG WANG
; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
; TITLE OF INVENTION: HORMONE ACTION IN PLANTS
; FILE REFERENCE: SALKINS. 046A
; CURRENT APPLICATION NUMBER: US/09/995.938A
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 335
; TYPE: PRT
; ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-8

Query Match 86.1%; Score 1519.5; DB 12; Length 335;
Best Local Similarity 88.4%; Pred. No. 2.1e-114;
Matches 297; Conservative 10; Mismatches 24; Indels 5; Gaps 4;

Qy 1 MTSFGATSTSAIAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGIAOGDYNLP 60
Db 1 MTSFGATSTSAIAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGIAOGDYNLP 59

Qy 61 KHCDDNEVLKALCYEAGWVVEEDGTTTKGCKPLPGELAGTSSRVTPYSSQNSPLSSAF 120
Db 61 KHCDDNEVLKALCYEAGWVVEEDGTTTKGCKPLPGELAGTSSRVTPYSSQNSPLSSAF 119

Qy 121 QSPISYQVSPSSSSSPSPSRGEPNNNSSTFFPFLRNGIIPSSLSLRISNCPVTPPV 179
Db 120 DSPISYQVSPSSSSSPSPSRGEPNNNSSTFFPFLRNGIIPSSLSLRISNCPVTPPV 177

Qy 180 VSSPTSCKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIPEC 238

Db 178 VSSPTSRNPKPLPTWESFTKQSMMAAKQSMSTSLNVPFAVAAPASPTHRRQPHATIP 237
Oy 239 ECDSDSSTVSGHWISFOKFAQQOOPFSASMTPTSTFNLVKPAPQMSPTAAFOELIQ 298
Db 238 ECDSDSSTVSGHWISFOKFAQQOOPFSASMTPTSTFNLVKPAPQMSPTAAFOELIQ 297
Oy 299 SSEPKFENSQVMPGGERIHQVGMEDLETLGNGKA 334
Db 298 SSEPKFENSQVMPGGERIHQVGMEDLETLGNGKA 333

RESULT 5

US-09-940-746-20
Sequence 20, Application US/09840746
Publication No. US20030166501A1
GENERAL INFORMATION:
APPLICANT: Chen, Huel-Mei
APPLICANT: Honchell, Cynthia D.
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Nuclein-Related Tumor Marker
FILE REFERENCE: PC-0039 US
CURRENT APPLICATION NUMBER: US/09/840.746
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 20
LENGTH: 528
TYPE: PRT
ORGANISM: Sus scrofa
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Genbank ID No. US20030166501A1 9915208
US-09-840-746-20

Query Match 7.8%; Score 137.5; DB 12; Length 528;
Best Local Similarity 24.6%; Pred. No. 0.0085;
Matches 61; Conservative 32; Mismatches 88; Indels 67; Gaps 9;

Oy 93 PLPG--EIACTSSRVTPYSSONOSPLSAPQSPIS-----YQVPS 133
Db 237 PIPSTTSVQPSSSSAPTATSVQSSSSTPIPTTSVQPSSSSAPTATSVQPS 296
Oy 134 SSP-----SPSRGEPPNNMSSTFPPLRNGIP--SLPLRLINSCEVTPVSS-- 182
Db 297 SSPPISSTTSVQPSSSSPTSTTSVQPSSSGAPTATSVQPSSS--SPISSTI 354
Oy 183 ---PTSKNPKPLPNWESIAKQSMASFPYVAVASAPSPHRRQPHATIP 239
Db 355 SVQPSSSSPTSTTSVQPSSSGAPTATSVQPSSSSVPT-----TSATSVR 405
Oy 240 CDESSTVDSGHWISFOKFAQQOOPFSASMTPTSTFNLVKPAPQ 284
Db 406 SSSSSSTPIPT-----TTSVQPSSSSVPTTATSVQPSSSSTPIPTTSVQPS 457
Oy 285 QMSPTAA 292
Db 458 SSAPTSA 465

RESULT 6

US-10-043-487-311
Sequence 311, Application US/10043487
Publication No. US20030055220A1
GENERAL INFORMATION:
APPLICANT: HYBRIDGENICS
APPLICANT: Pierre, LEGRAIN
TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptid
FILE REFERENCE: B4778A
CURRENT APPLICATION NUMBER: US/10/043.487
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/261,130

PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 561
SOFTWARE: PatentIn version 3.1
SEQ ID NO 311
LENGTH: 314
TYPE: PRT
ORGANISM: Shigella flexneri
US-10-043-487-311

Query Match 7.4%; Score 131; DB 15; Length 314;
Best Local Similarity 29.7%; Pred. No. 0.015;
Matches 65; Conservative 23; Mismatches 91; Indels 40; Gaps 12;

Oy 93 PLPGEIACTSSRVTPYSSONOSPLSAPQSPISYQVPSSSSF--PSRGEPPNNMSST 151
Db 108 PTPSPSPSPNYP--TSPNVSPTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 161
Oy 152 FPPPLRNGCIPSSLPGLRLINSCEVTPPVSSPTSNPKP-----LPNWSIA-KQSM 202
Db 162 YTP-----SSPSYSPS--SPSYSPSPKYTPSPSYSPSPSPSPSPSPSPSPSPSPSP 213
Oy 203 AIAKQSMASFPY-----PFAVASAPASPTHRRQPHATIP 250
Db 214 TSPKYSPTSP 273
Oy 251 GHWISFOKFAQQOOPFSASMTPTSTFNLVKPAPQMSPT 289
Db 274 -YSPTPKSGSTYSPTSPKYSPTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 308

RESULT 7

US-10-105-004-2
Sequence 2, Application US/10105004
Publication No. US20030105002A1

GENERAL INFORMATION:

APPLICANT: Murray, Jeffrey

APPLICANT: Semina, Elena

TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR

NUMBER OF SEQUENCES: 139

CORRESPONDENCE ADDRESS: HOAG & ELIOT LLP

ADDRESS: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/105,004

FILING DATE: 22-Mar-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/754,477

FILING DATE: 22-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Arnold, Beth E.

REGISTRATION NUMBER: 35,430

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-7000

TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 271 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-105-004-2

Query Match 7.2%; Score 127.5; DB 15; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.023;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTGAAAAAARRRPSWRENNRRRRRAVAAKITGTGAAOGDYNLPKHCN 65
DB 73 AVMTNLTBARVWFKNRKAKRKRRNQALCKNGGPGQ-FNGIMOPYDDMYPGYSYN 131
QY 66 NEVLKALCVENAGWVEEDGTTTRKCKPLPGEIAGTSRVT-P-YSSONOSPIS--SAFOS 122
DB 132 NMAAKGL-----TSASLSTKSPFPFNSMNVNPLSSQSMFSP 167
QY 123 PIPSYOVSPSSSFPSPRGEPN-----NMSSTFPPLRNGIPPSLSLRISNCPV 176
DB 168 PMSISMSMSSMWPASVATGVPSSLSNLTNNLNLSPSL-NSAVPTP-----ACPY 218
QY 177 TPVVSFTSKNPPPLPWESIAKOSMAI-----AKOSMASFNYPYAVSAPAS 224
DB 219 APPT-----PPVYRDTCNSSLASLRKAKQH--SFGYA--SVOKPAS 258

RESULT 8
US-10-105-004-121
; Sequence 121, Application US/10105004
; Publication No. US20030105002A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Jeffrey
; TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
; AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 139
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HONG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/105,004
; FILING DATE: 22-Mar-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/754,477
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIA-022.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 121:
US-10-105-004-121

Query Match 7.2%; Score 127.5; DB 15; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.023;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTGAAAAAARRRPSWRENNRRRRRAVAAKITGTGAAOGDYNLPKHCN 65
DB 73 AVMTNLTBARVWFKNRKAKRKRRNQALCKNGGPGQ-FNGIMOPYDDMYPGYSYN 131

QY 66 NEVLKALCVENAGWVEEDGTTTRKCKPLPGEIAGTSRVT-P-YSSONOSPIS--SAFOS 122
DB 132 NMAAKGL-----TSASLSTKSPFPFNSMNVNPLSSQSMFSP 167
QY 123 PIPSYOVSPSSSFPSPRGEPN-----NMSSTFPPLRNGIPPSLSLRISNCPV 176
DB 168 PMSISMSMSSMWPASVATGVPSSLSNLTNNLNLSPSL-NSAVPTP-----ACPY 218
QY 177 TPVVSFTSKNPPPLPWESIAKOSMAI-----AKOSMASFNYPYAVSAPAS 224
DB 219 APPT-----PPVYRDTCNSSLASLRKAKQH--SFGYA--SVOKPAS 258

RESULT 9
US-10-105-004-125
; Sequence 125, Application US/10105004
; Publication No. US20030105002A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Jeffrey
; TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
; AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 139
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HONG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/105,004
; FILING DATE: 22-Mar-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/754,477
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIA-022.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 125:
US-10-105-004-125

Query Match 7.2%; Score 127.5; DB 15; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.023;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTGAAAAAARRRPSWRENNRRRRRAVAAKITGTGAAOGDYNLPKHCN 65
DB 73 AVMTNLTBARVWFKNRKAKRKRRNQALCKNGGPGQ-FNGIMOPYDDMYPGYSYN 131
QY 66 NEVLKALCVENAGWVEEDGTTTRKCKPLPGEIAGTSRVT-P-YSSONOSPIS--SAFOS 122
DB 132 NMAAKGL-----TSASLSTKSPFPFNSMNVNPLSSQSMFSP 167
QY 123 PIPSYOVSPSSSFPSPRGEPN-----NMSSTFPPLRNGIPPSLSLRISNCPV 176
DB 168 PMSISMSMSSMWPASVATGVPSSLSNLTNNLNLSPSL-NSAVPTP-----ACPY 218

Qy 177 TPVSSPTSKNPKPLNPMWESIAKQSMAL-----AKQSMASPNYPFAVASAPAS 224
 Db 219 APPT-----PPVIVYDTCNSSLASLRKAKQH--SSFGVA--SVQKFPAS 258

RESULT 10
 US-09-858-754-4
 ; Sequence 4, Application US/09858754
 ; Patent No. US20020055130A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Johnson, Gary L.
 ; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING APOPTOSIS
 ; FILE REFERENCE: CPI-042
 ; CURRENT APPLICATION NUMBER: US/09/858,754
 ; CURRENT FILING DATE: 2001-05-16
 ; PRIOR APPLICATION NUMBER: 09/023,130
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: 60/039,740
 ; PRIOR FILING DATE: 1997-02-14
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 1493
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; US-09-858-754-4

Query Match 7.2%; Score 127.5; DB 9; Length 1493;
 Best Local Similarity 22.2%; Pred. No. 0.2;
 Matches 76; Conservative 38; Mismatches 124; Indels 99; Gaps 12;

Qy 5 GATSTSAALAAAAAARRRKPSWRENNRRRRRAVAAKIYTGRLAOGDYNLPKHCD 64
 Db 134 GATSPAGAEPPSAA-----PSGREMENK-----ETLKGLHKMD--RPERRM 174
 Qy 65 NNEVLKALCVAEAGVVEEDGTTTKG---CKPLPEIAGTSSRTTPYSQONSPISAFQ 121
 Db 175 IRELUKATCMHA-WKHEKLERNRNRGPVVVPRIP--IKGDSEMSNLAELQEGQACSA 231
 Qy 122 SPIPSYVSSPSSSPSPSRGEPNNNSSTFPFLRNGIPSLPSLRISNCPV----- 176
 Db 232 APAVKGRSPSPGSSPGRSGKPE-----PGVARKVSPVPPGSG 272
 Qy 177 --TPVSSPTSKNPKPLNPMWESIAKQSMASPNYPFAVASAPAPTR----- 228
 Db 273 RITPPRAAPDPDGFSPYSPEBTSRRVKNKRAQLYLQOIGPNSFLLGSDPDKKRYFI 332
 Qy 229 -----HGHITLTIPECDESD-----SSTVDS 250
 Db 333 GPONCGGCGTFCIHLFLVMLRVFQLEPSDPMWRKTLKNPEVESLPQKHSRRSRKA 392
 Qy 251 GHWISFOKFAQO-----QPFASMWPTSPFTNLVAPAOQMS 288
 Db 393 PSRNTIQKFPVSRMSNCHTSSSSSTSTSSSENSIKDBEEQWCP 434

RESULT 11
 US-10-106-698-5571
 ; Sequence 5571, Application US/10106698
 ; Publication No. US20030109630A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruden et al.
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
 ; FILE REFERENCE: PA005P1
 ; CURRENT APPLICATION NUMBER: US/10/106,698
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US 60/157,137
 ; PRIOR FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: US 60/163,280
 ; PRIOR FILING DATE: 1999-11-03

Qy 177 TPVSSPTSKNPKPLNPMWESIAKQSMAL-----AKQSMASPNYPFAVASAPAS 224
 Db 219 APPT-----PPVIVYDTCNSSLASLRKAKQH--SSFGVA--SVQKFPAS 258

NUMBER OF SEQ ID NOS: 9564
 ; SOFTWARE: PatentIn Ver. 3.0
 ; SEQ ID NO 5571
 ; LENGTH: 210
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (137)
 ; OTHER INFORMATION: Xaa
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (142)
 ; OTHER INFORMATION: Xaa
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (171)
 ; OTHER INFORMATION: Xaa
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (179)
 ; OTHER INFORMATION: Xaa
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (193)
 ; OTHER INFORMATION: Xaa
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (197)
 ; OTHER INFORMATION: Xaa
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (210)
 ; OTHER INFORMATION: Xaa
 ; US-10-106-698-5571

Query Match 7.2%; Score 126.5; DB 15; Length 210;
 Best Local Similarity 25.8%; Pred. No. 0.02;
 Matches 48; Conservative 21; Mismatches 74; Indels 43; Gaps 7;

Qy 6 ATSTSAALAAAAAARRRKPSWRENNRRRRRAVAAKIYTGRLAOGDYNLPKHCD 65
 Db 38 AWTNLTEARVRYVFKNRAKRRERNOQABLCKNGFGPO-FNCLMOPYDMDYPOYSYN 96
 Qy 66 NEVLKALCVAEAGVVEEDGTTTKG---CKPLPEIAGTSSRTTPYSQONSPISAFQ 121
 Db 97 NMAKGL-----TSASLSTKSPFPFPMVMVNPSSOSMFS 132
 Qy 123 PIPEYQVSPSSSPSPSRGEPN-----NNMSSTFPFLRNGIPSLPSLRISNCPV 176
 Db 133 PNSTXSMSSMSSNPVAVTGVPSSLSLNLNLNLSSPXL-NSAVPTX-----ACPY 183
 Qy 177 TPVSS 182
 Db 184 APPTSS 189

RESULT 12
 US-10-105-004-5
 ; Sequence 5, Application US/10105004
 ; Publication No. US20030105002A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Murray, Jeffrey
 ; TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
 ; AND DIAGNOSTIC USES THIREROR
 ; NUMBER OF SEQUENCES: 139
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

Qy 177 TPVSSPTSKNPKPLNPMWESIAKQSMAL-----AKQSMASPNYPFAVASAPAS 224
 Db 219 APPT-----PPVIVYDTCNSSLASLRKAKQH--SSFGVA--SVQKFPAS 258

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/105,004
FILING DATE: 22-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UTA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-105-004-5

Query Match
Best Local Similarity 25.9%; Pred. No. 0.028; Length 271;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

6 ATSTSAATAAAARRRRRRRRRRRRAAATYTLRAQGDYLNPKHCDN 65
73 AVWTNLTLEAVRWVFKRRAKRRERNOQALCKNGFGPO-FNGLMQPYDMYGYSTN 131
66 NEVLKALCVAAGWVEEDGTTYRKCKPLGEIAGTSSRVTP-VSSONQSPLS--SAFOS 122
132 NMAKGL-----TSASLSTKSPFFFSNMVNPFLSSQSMFSP 167
123 PISYQVSPSSSPSPSPSGEPN-----NNMSTFFPLRNGGIPPSLPRLRISNCPV 176
168 PMSISSMSMSSMNPVSAVTGVPGLSLNLSLNNLNLSPSL-NSAVPTP-----ACPY 218
177 TPVSSPFTSKPKPLPNMESIAKQSMAL---AKQSMASFNPYPVAVSAPAS 224
219 APPT-----PPVYRDTCNSSLASLRKAKQH--SSFCYA--SYQNPAAS 258

RESULT 13
US-10-032-585-7168
Sequence 7168, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jang
APPLICANT: Charles, Boone
APPLICANT: Howard, Buseasy
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7168
LENGTH: 1265
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7168

Query Match
Best Local Similarity 7.2%; Score 126.5; DB 12; Length 1265;
Matches 58; Conservative 34; Mismatches 89; Indels 59; Gaps 9;

93 PLPEIAGTSSRVTP-PVS-----SQNQSPLSAFQSPISYQVSPSSSPSPSGRGEPPN 146
965 PPSISISSTRTTTPSTVPAKPTPVSAKVAATVTSQSPSSGSPSTLNSPR 1024
147 NMSSTFFPLRNGGIP-----SSLPLRLISNCGVTPPVVSSPFTSKPKPLPNMESIAKQSM 202

1025 LAKNPYAPSVTBEQLPKPKISYATPEPAHLNNGSPSTSVAPP--KNPYAVPSTSVSHAGI 1082
203 AIAKQSMASFNVPFVAVSAPASTHRQFHTLATIPEDCSDSDSTVSGHMISQKFAQO 262
1083 APPPPA-----PKQSAAPP----- 1099
263 QPFSASG-VPTSPFTPLVVKAPQOM--SPPTAFAOEIGQSSSEFKFE-NSQVKEPGEERIH 318
1100 QPFGSSMSMNPVDAFNGVPPPPVPGRAVSTPAALKEQPPAREPELPVOSKHKGRTH 1159

RESULT 14
US-10-105-004-131
Sequence 131, Application US/10105004
Publication No. US20030105002A1
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
Semina, Elena
TITLE OF INVENTION: REG COMPOSITIONS AND THERAPEUTIC
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/105,004
FILING DATE: 22-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UTA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 131:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 131:
US-10-105-004-131

Query Match
Best Local Similarity 7.1%; Score 124.5; DB 15; Length 271;
Matches 60; Conservative 24; Mismatches 89; Indels 59; Gaps 11;

6 ATSTSAATAAAARRRRRRRRRRRRAAATYTLRAQGDYLNPKHCDN 65
73 AVWTNLTLEAVRWVFKRRAKRRERNOQALCKNGFGPO-FNGLMQPYDMYGYSTN 131
66 NEVLKALCVAAGWVEEDGTTYRKCKPLPEIAGTSSRVTP-VSSONQSPLS--SAFOS 122
132 NMAKGL-----TSASLSTKSPFFFSNMVNPFLSSQSMFSP 167
123 PISYQVSPSSSPSPSPSGEPN-----NNMSTFFPLRNGGIPPSLPRLRISNCPV 176
168 PMSISSMSMSSMNPVSAVTGVPGLSLNLSLNNLNLSPSL-NSAVPTP-----ACPY 218
177 TPVSSPFTSKPKPLPNMESIAKQSMAL---AKQSMASFNPYPVAVSAPAS 224

Db 219 APPT-----PPYVYRDTCNSSLSALRLKAKQH-SSFVGA--SVQKPS 258

RESULT 15

US-10-294-804-2
 / Sequence 2, Application US/10294804
 / Publication No. US20030133948A1
 / GENERAL INFORMATION:
 / APPLICANT: Robertson, Etle S.
 / APPLICANT: Coctet, Murray A.
 / TITLE OF INVENTION: Methods to inhibit or enhance the binding of viral DNA
 / FILE REFERENCE: UM-03778
 / CURRENT APPLICATION NUMBER: US/10/294,804
 / PRIOR FILING DATE: 2002-11-14
 / PRIOR APPLICATION NUMBER: US/09/410,399
 / NUMBER OF SEQ ID NOS: 6
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 2
 / LENGTH: 1162
 / TYPE: PRT
 / ORGANISM: Kaposi's sarcoma-associated herpesvirus
 / US-10-294-804-2

Query Match 7.0%; Score 124; DB 12; Length 1162;
 Best Local Similarity 22.5%; Pred. No. 0.28;
 Matches 79; Conservative 49; Mismatches 129; Indels 94; Gaps 17;

QY 8 STSAAAAAAAAARRKPSNRERENRRRRRAVAAKIYTGLRAGDYNLPKCDNNE 67
 DB 13 STCAPLTRGSCRKRRKSPERCDLGDLDLHLPKRKHVADSI--DRECGPHTLP----- 63
 QY 68 VLKALCVAGWVVEEDGTYRKGCKPLGEIAGTSRVTPTYSONOS-----PLSS 118
 DB 64 -----IPGSPTVFTSG--LPAFVSPTLPVAPISPAPATPLPPALLPVT 108
 QY 119 AFQSPF-PSYQVSSSSSPSPSRG-EPNNMSTFPFLRNGIPSSLPLRLISNCPV 176
 DB 109 TSSSPIPSPHVSFGTDTHTSPSPALPPTQSPSSORP-----PLSPYGRPDSSTPM 161
 QY 177 TPVVS-----SPTSKNPKP--LPNVESTAKQSM-AIAQSMASFVYPF---YATSA 222
 DB 162 RPPPSQOTTTPPHFTTTPPEPPSKSPDSLAPSTLRSLRKRLSSPQSPSTLNPICOSP 221
 QY 223 ASPTHRHQFTLATIPECDSDSSTVDGHWISFOKEAQQPF---SAS-----MVPT 272
 DB 222 VSP-----PRCDFANRNV-----YPPWATESPIYVSSSDGDTPPRQPT 261
 QY 273 SP-TFNLVKPAPQOMSPNTAF-----OEIGSSSEKPKENSQVKPWEGB 315
 DB 262 SPISIGSSSPSGWGDITAMLVLLAIEASKNKEKSENNQAGEDNGD 312

Search completed: December 19, 2003, 14:32:05
 Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2003, 14:24:08 ; Search time 43 Seconds
(without alignments)
1240.282 Million cell updates/sec

Title: US-09-995-938a-7

Perfect score: 1765
Sequence: 1 MTSDGATSTSAIAAAAAAAAAA.....IHDVGMEDLELTGKXKRG 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_19JUN03:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1758	99.6	21	AA16755
2	1511.5	85.6	21	AA16755
3	1511.5	85.6	21	AA16755
4	1447	82.0	21	AA16755
5	1447	82.0	21	AA16755
6	1052	59.6	21	AA16755
7	1052	59.6	21	AA16755
8	994	56.3	21	AA16755
9	541	30.7	21	AA16755

10	541	30.7	284	21	AA16755	Arabidopsis thaliana
11	330	18.7	169	21	AA16755	Arabidopsis thaliana
12	235.5	13.3	205	21	AA16755	Arabidopsis thaliana
13	235.5	13.3	205	21	AA16755	Arabidopsis thaliana
14	235.5	13.3	205	21	AA16755	Arabidopsis thaliana
15	235.5	13.3	205	21	AA16755	Arabidopsis thaliana
16	203.5	11.5	687	23	AA16755	Arabidopsis thaliana
17	190	10.8	689	23	AA16755	Arabidopsis thaliana
18	131	7.4	314	23	AA16755	Arabidopsis thaliana
19	127	7.2	314	23	AA16755	Arabidopsis thaliana
20	126.5	7.2	210	22	AA16755	Arabidopsis thaliana
21	126.5	7.2	1265	22	AA16755	Arabidopsis thaliana
22	126	7.1	904	22	AA16755	Arabidopsis thaliana
23	124.5	7.1	904	22	AA16755	Arabidopsis thaliana
24	124.5	7.1	1887	22	AA16755	Arabidopsis thaliana
25	124	7.0	1162	21	AA16755	Arabidopsis thaliana
26	124	7.0	1162	21	AA16755	Arabidopsis thaliana
27	124	7.0	1162	22	AA16755	Arabidopsis thaliana
28	124	7.0	1162	22	AA16755	Arabidopsis thaliana
29	123.5	7.0	259	21	AA16755	Arabidopsis thaliana
30	123.5	7.0	268	21	AA16755	Arabidopsis thaliana
31	123	7.0	989	21	AA16755	Arabidopsis thaliana
32	123	7.0	994	21	AA16755	Arabidopsis thaliana
33	123	7.0	1006	21	AA16755	Arabidopsis thaliana
34	122.5	6.9	1549	22	AA16755	Arabidopsis thaliana
35	122.5	6.9	2414	16	AA16755	Arabidopsis thaliana
36	122.5	6.9	2414	19	AA16755	Arabidopsis thaliana
37	122.5	6.9	2414	23	AA16755	Arabidopsis thaliana
38	122.5	6.9	2414	24	AA16755	Arabidopsis thaliana
39	122.5	6.9	2414	24	AA16755	Arabidopsis thaliana
40	122.5	6.9	2414	24	AA16755	Arabidopsis thaliana
41	122.5	6.9	2414	24	AA16755	Arabidopsis thaliana
42	122.5	6.9	2414	24	AA16755	Arabidopsis thaliana
43	122	6.9	1228	24	AA16755	Arabidopsis thaliana
44	122	6.9	1475	22	AA16755	Arabidopsis thaliana
45	122	6.9	3238	22	AA16755	Arabidopsis thaliana

ALIGNMENTS

RESULT 1
AA16755
ID AA16755 standard; Protein; 336 AA.
AC AA16755;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17522.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0121810.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140883.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142350.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145115.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158359.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 85.6%; Score 1511.5; DB 21; Length 335;
Best Local Similarity 88.1%; Pred. No. 2, 2e-121;
Matches 296; Conservative 10; Mismatches 25; Indels 5; Gaps 4;

QY 1 MTSDCATSTSAATAAARRRKPSWRERENRRERRRAVAAKIYGLRAQGVNLP 60
DB 1 MTSDCATSTS-AAAAAATAATRRKPSWRERENRRERRRAVAAKIYGLRAQGVNLP 59
QY KCDNNEVLKALCEVAGWVEEDGTYRKCKPLPEIAGTSSRVTPYSSONOSPSSAF 120
DB KCDNNEVLKALCEVAGWVEEDGTYRKCKPLPEIAGTSSRVTPYSSONOSPSSAF 119
QY 121 QSPIPSYOVSPSSSPSPSPSR-GEPNNNMSTFPFLRNGCIPSSLPRLISNSCPTTP 179
DB 120 DSPILSYOVSPSSSPSPSPSRVGDPHN--ISTIFFLRNGCIPSSLPRLISNSAPVTPP 177
QY 180 VSPFTSKNPKPLPNWMSIAKOSMAI-AKOSMAFNYPFVAVSAPASTHROHPTLATIP 238
DB 178 VSPFTSKNPKPLPNWMSIAKOSMAI-AKOSMAFNYPFVAVSAPASTHROHPTLATIP 237
QY 239 ECDSDSSTVDSGHWISFOKFAQQOPFSASMYTSTFNLVKPAPQOMSPTNAFQBIQ 298
DB 238 ECDSDSSTVDSGHWISFOKFAQQOPFSASMYTSTFNLVKPAPQOLSPTNAFQBIQ 297
QY 299 SSEFKFENGQVKNPGEGRHIDVAMEDELTGNGKA 334
DB 298 SSEFKFENGQVKNPGEGRHIDVAMEDELTGNGKA 333
RESULT 3
AAG15884 standard; Protein; 335 AA.
ID AAG15884
XX AAG15884;
AC AAG15884;
XX


```
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 22-OCT-1999; 99US-0161405.
PR 23-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 85.6%; Score 1511.5; DB 21; Length 335;
Best Local Similarity 88.1%; Pred. No. 2.2e-121; Indels 5; Gaps 4;
Matches 296; Conservative 10; Mismatches 25;

```
QY 1 MTSDCATSTSAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGRLAQCQDYNLP 60
DB 1 MTSDCATSTSAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGRLAQCQDYNLP 59
QY 61 KHCNNNEVLKALCYEAGNVVEBDGTTTKKCKPLDGEIAGTSSRYTPYSSQNSQSLSSNAF 120
DB 60 KHCNNNEVLKALCYEAGNVVEBDGTTTKKCKPLDGEIAGTSSRYTPYSSQNSQSLSSNAF 119
QY 121 QSPITPSYVSPSSSSSPSPR-GEENNMSSTPFPFLNCGTSSLPRLISNSCPVTIP 179
DB 120 DSPITPSYVSPSSSSSPSPR-GEENNMSSTPFPFLNCGTSSLPRLISNSCPVTIP 177
QY 180 VSSPSTSKPKPLPNNESTIAKOSMAI-AKOSMASFYVPFYAVASAPAPTHRHQFHTLATIP 238
DB 178 VSSPSTSKPKPLPNNESTIAKOSMAI-AKOSMASFYVPFYAVASAPAPTHRHQFHTLATIP 237
QY 239 ECDSDSSTVDSGHWISFQKFAQQOPFSASMTPTSTFPLVLPAPQNSPNTAAIOEIQO 298
```

```
DB 238 ECDSDSSTVDSGHWISFQKFAQQOPFSASMTPTSTFPLVLPAPQNSPNTAAIOEIQO 297
QY 299 SSEFKFENSOVKPWEGEIRIHDCMEDLELTIGSKA 334
DB 298 SSEFKFENSOVKPWEGEIRIHDCMEDLELTIGSKA 333

RESULT 4
AAG14238
ID AAG14238 standard; Protein; 318 AA.
AC AAG14238;
XX
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14023.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0901439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
```

PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139464.
 PR 21-JUN-1999; 99US-0139465.
 PR 21-JUN-1999; 99US-0139466.
 PR 22-JUN-1999; 99US-0139467.
 PR 23-JUN-1999; 99US-0139468.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140823.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142803.
 PR 12-JUL-1999; 99US-0142920.
 PR 13-JUL-1999; 99US-0142927.
 PR 14-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144086.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144332.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 23-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146388.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149375.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 26-AUG-1999; 99US-0150566.
 PR 27-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151438.
 PR 07-SEP-1999; 99US-0151930.
 PR 10-SEP-1999; 99US-0152363.
 PR 13-SEP-1999; 99US-0153070.
 PR 15-SEP-1999; 99US-0153758.
 PR 16-SEP-1999; 99US-0154018.
 PR 20-SEP-1999; 99US-0154039.
 PR 22-SEP-1999; 99US-0154779.
 PR 23-SEP-1999; 99US-0155139.
 PR 24-SEP-1999; 99US-0155466.
 PR 28-SEP-1999; 99US-0155659.
 PR 29-SEP-1999; 99US-0156458.
 PR 04-OCT-1999; 99US-0156596.
 PR 05-OCT-1999; 99US-0157717.
 PR 06-OCT-1999; 99US-0157753.
 PR 07-OCT-1999; 99US-0157865.
 PR 08-OCT-1999; 99US-0158029.
 PR 12-OCT-1999; 99US-0158232.
 PR 13-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 14-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 21-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 25-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 28-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 29-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 82.0%; Score 1447; DB 21; Length 318;
 Best Local Similarity 88.0%; Pred. No. 7, 1e-116;
 Matches 279; Conservative 10; Mismatches 24; Indels 4; Gaps 3;

QY 20 AARRKPSWRENNRRRRRAVAAKITYTGLRAQGDVNLPGHCDNNEVYKALCYEAGWV 79
 DB 2 ATRRRKPSWRENNRRRRRAVAAKITYTGLRAQGNVNLPRGCDNNEVYKALCSEAGWV 61
 QY 80 VEDGTTYRKCKPLPGELIAGTSSRVTFYSSQNSPLSSAFQSPYQVSPSSSFSPSP 139

DB 62 VERDGTTRKGRHPLGDMAGSSSRATPTSSHNQSPUSSTFDSPILISTYQVSPSSSSFPSP 121
QY 140 SR-GEPPNNMSTFPPLRNGGIPSSLPLSLRISNSCVPVPVSPSKPKPLPNWESIA 198
DB 122 SRGDDPHN--ISITFPPLRNGGIPSSLPLRISNSAVPTFPVSPSRNSKPLPTWESFT 179
QY 199 KQSMAL-AKQSMASFNYPFYAVASAPASPTRRHQPHTLATIPECDSDSSSTVDSGHWISFQ 257
DB 180 KQSMMAAKQSMSTSLNYPFYAVASAPASPTRRHQPHTLATIPECDSDSSSTVDSGHWISFQ 239
QY 258 KFAQOQPFPSASWPTSTPTFNLVVKPAPQOMSPTAAFOEIQSSSEPKENSQVKEGGERI 317
DB 240 KFAQOQPFPSASWPTSTPTFNLVVKPAPQOLSPTTAIOEIQSSSEPKENSQVKEGGERI 239
QY 318 HDVGMEDLETLGNGKA 334
DB 300 HDVAMEDLETLGNGKA 316

RESULT 5
AAG15885
ID AAG15885 standard; Protein: 318 AA..
AC AAG15885;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 16312.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; .99US-0121825.
XX 05-MAR-1999; .99US-0123180.
XX 09-MAR-1999; .99US-0123548.
XX 23-MAR-1999; .99US-0125788.
XX 23-MAR-1999; .99US-0126264.
XX 29-MAR-1999; .99US-0126785.
XX 01-APR-1999; .99US-0127462.
XX 06-APR-1999; .99US-0128234.
XX 08-APR-1999; .99US-0128714.
XX 16-APR-1999; .99US-0128845.
XX 19-APR-1999; .99US-0130077.
XX 21-APR-1999; .99US-0130449.
XX 23-APR-1999; .99US-0130510.
XX 23-APR-1999; .99US-0130891.
XX 28-APR-1999; .99US-0131449.
XX 30-APR-1999; .99US-0132048.
XX 30-APR-1999; .99US-0132407.
XX 04-MAY-1999; .99US-0132484.
XX 05-MAY-1999; .99US-0132485.
XX 06-MAY-1999; .99US-0132486.
XX 06-MAY-1999; .99US-0132487.
XX 07-MAY-1999; .99US-0132863.
XX 11-MAY-1999; .99US-0134256.
XX 14-MAY-1999; .99US-0134218.
XX 14-MAY-1999; .99US-0134219.
XX 14-MAY-1999; .99US-0134221.
XX 14-MAY-1999; .99US-0134370.
XX 18-MAY-1999; .99US-0134768.
XX 19-MAY-1999; .99US-0134941.
XX 20-MAY-1999; .99US-0135124.
XX 21-MAY-1999; .99US-0135153.

PR 24-MAY-1999; .99US-0135629.
PR 25-MAY-1999; .99US-0136021.
PR 27-MAY-1999; .99US-0136392.
PR 28-MAY-1999; .99US-0136782.
PR 01-JUN-1999; .99US-0137222.
PR 03-JUN-1999; .99US-0137528.
PR 04-JUN-1999; .99US-0137502.
PR 07-JUN-1999; .99US-0137724.
PR 08-JUN-1999; .99US-0138094.
PR 10-JUN-1999; .99US-0138540.
PR 10-JUN-1999; .99US-0138847.
PR 14-JUN-1999; .99US-0139119.
PR 16-JUN-1999; .99US-0139452.
PR 16-JUN-1999; .99US-0139453.
PR 17-JUN-1999; .99US-0139492.
PR 18-JUN-1999; .99US-0139454.
PR 18-JUN-1999; .99US-0139455.
PR 18-JUN-1999; .99US-0139456.
PR 18-JUN-1999; .99US-0139457.
PR 18-JUN-1999; .99US-0139458.
PR 18-JUN-1999; .99US-0139459.
PR 18-JUN-1999; .99US-0139460.
PR 18-JUN-1999; .99US-0139461.
PR 18-JUN-1999; .99US-0139462.
PR 18-JUN-1999; .99US-0139463.
PR 18-JUN-1999; .99US-0139750.
PR 18-JUN-1999; .99US-0139753.
PR 21-JUN-1999; .99US-0139817.
PR 22-JUN-1999; .99US-0139899.
PR 23-JUN-1999; .99US-0140353.
PR 23-JUN-1999; .99US-0140354.
PR 24-JUN-1999; .99US-0140695.
PR 28-JUN-1999; .99US-0140823.
PR 29-JUN-1999; .99US-0140991.
PR 30-JUN-1999; .99US-0141287.
PR 01-JUL-1999; .99US-0141842.
PR 01-JUL-1999; .99US-0142154.
PR 02-JUL-1999; .99US-0142305.
PR 06-JUL-1999; .99US-0142303.
PR 08-JUL-1999; .99US-0142803.
PR 09-JUL-1999; .99US-0142920.
PR 12-JUL-1999; .99US-0142977.
PR 13-JUL-1999; .99US-0143542.
PR 14-JUL-1999; .99US-0143624.
PR 15-JUL-1999; .99US-0144005.
PR 16-JUL-1999; .99US-0144085.
PR 16-JUL-1999; .99US-0144086.
PR 19-JUL-1999; .99US-0144325.
PR 19-JUL-1999; .99US-0144331.
PR 19-JUL-1999; .99US-0144332.
PR 19-JUL-1999; .99US-0144333.
PR 19-JUL-1999; .99US-0144334.
PR 19-JUL-1999; .99US-0144335.
PR 20-JUL-1999; .99US-0144335.
PR 20-JUL-1999; .99US-0144632.
PR 20-JUL-1999; .99US-0144684.
PR 20-JUL-1999; .99US-0144814.
PR 21-JUL-1999; .99US-0145088.
PR 21-JUL-1999; .99US-0145088.
PR 22-JUL-1999; .99US-0145085.
PR 22-JUL-1999; .99US-0145087.
PR 22-JUL-1999; .99US-0145089.
PR 22-JUL-1999; .99US-0145192.
PR 22-JUL-1999; .99US-0145192.
PR 23-JUL-1999; .99US-0145218.
PR 23-JUL-1999; .99US-0145224.
PR 26-JUL-1999; .99US-0145276.
PR 27-JUL-1999; .99US-0145913.
PR 27-JUL-1999; .99US-0145918.
PR 27-JUL-1999; .99US-0145919.
PR 28-JUL-1999; .99US-0145951.
PR 02-AUG-1999; .99US-0146386.
PR 02-AUG-1999; .99US-0146386.

```

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

```

```

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Beet Local Similarity 82.0%; Score 1447; DB 21; Length 318;
Matches 279; Conservative 10; Mismatches 24; Indels 4; Gaps 3;

QY 20 AARRKPSWREERNRRRRRAVAKIYTGIRAGQDYNLPFGHCNNNEVLKALCIEAGNV 79
DB 2 ATRRKPSWREERNRRRRRAVAKIYTGIRAGQDYNLPFGHCNNNEVLKALCIEAGNV 61
QY 80 VEEDGTYRKCKPLPGEIAGTSRVTPYSSONOSPILSAFOSPISYOVSSSSSPSP 139
DB 62 VEEDGTYRKCKPLPGEIAGTSRVTPYSSONOSPILSAFOSPISYOVSSSSSPSP 121
QY 140 SR-CEPNNNSSTFPPLRNGCIPSSLPBLRISNCPVTPPVSSPTSKPKPLPNNESIA 198
DB 122 SRVGDPHN--ISTLPPLRNGCIPSSLPBLRISNCPVTPPVSSPTSKPKPLPNNESIA 179
QY 199 KQSMAL-AKQSMASFNYPFVAVSAPASPTAHQHTLATITECDSDSDSTVDSGHWISFO 257
DB 180 KQSMALAKQSMASFNYPFVAVSAPASPTAHQHTLATITECDSDSDSTVDSGHWISFO 239
QY 258 KFAQQCPFSASMPYPTSPFLVLPKAPQOMSPNTAFOEIGQSSPKFENSQVKEGRI 317
DB 240 KFAQQCPFSASMPYPTSPFLVLPKAPQOLSPNTAAOEGIGQSSPKFENSQVKEGRI 299
QY 318 HDVGMEDLETLANGKA 334
DB 300 HDVAMEDLETLANGKA 316

RESULT 6
AAAG14239
ID AAAG14239 standard; Protein; 239 AA.
XX AC
XX AAAG14239;
XX DT
XX 17-OCT-2000 (first entry)
XX DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 14024.
XX KW
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW
XX Arabidopsis thaliana.
XX OS
XX Arabidopsis thaliana.
XX PN
XX EP1033405-A2.
XX PD
XX 06-SEP-2000.
XX PF
XX 25-FEB-2000; 2000EP-0301439.
XX XX
XX 25-FEB-1999; 99US-01231825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.

```

PR	05-MAY-1999;	99US-0132485;	PR	22-JUL-1999;	99US-0145087;
PR	06-MAY-1999;	99US-0132486;	PR	22-JUL-1999;	99US-0145089;
PR	06-MAY-1999;	99US-0132487;	PR	22-JUL-1999;	99US-0145192;
PR	07-MAY-1999;	99US-0132863;	PR	23-JUL-1999;	99US-0145145;
PR	11-MAY-1999;	99US-0132856;	PR	23-JUL-1999;	99US-0145218;
PR	14-MAY-1999;	99US-0134218;	PR	23-JUL-1999;	99US-0145224;
PR	14-MAY-1999;	99US-0134219;	PR	26-JUL-1999;	99US-0145276;
PR	14-MAY-1999;	99US-0134221;	PR	27-JUL-1999;	99US-0145913;
PR	14-MAY-1999;	99US-0134370;	PR	27-JUL-1999;	99US-0145919;
PR	18-MAY-1999;	99US-0134376;	PR	28-JUL-1999;	99US-0145931;
PR	19-MAY-1999;	99US-0134941;	PR	02-AUG-1999;	99US-0146386;
PR	20-MAY-1999;	99US-0135124;	PR	02-AUG-1999;	99US-0146388;
PR	21-MAY-1999;	99US-0135353;	PR	02-AUG-1999;	99US-0147038;
PR	24-MAY-1999;	99US-0136629;	PR	03-AUG-1999;	99US-0147204;
PR	25-MAY-1999;	99US-0136021;	PR	04-AUG-1999;	99US-0147302;
PR	28-MAY-1999;	99US-0136392;	PR	04-AUG-1999;	99US-0147307;
PR	28-MAY-1999;	99US-0136782;	PR	05-AUG-1999;	99US-0147132;
PR	01-JUN-1999;	99US-0137222;	PR	05-AUG-1999;	99US-0147260;
PR	03-JUN-1999;	99US-0137528;	PR	06-AUG-1999;	99US-0147303;
PR	04-JUN-1999;	99US-0137502;	PR	06-AUG-1999;	99US-0147416;
PR	07-JUN-1999;	99US-0137724;	PR	09-AUG-1999;	99US-0147493;
PR	08-JUN-1999;	99US-0138094;	PR	09-AUG-1999;	99US-0147935;
PR	10-JUN-1999;	99US-0138840;	PR	10-AUG-1999;	99US-0148171;
PR	10-JUN-1999;	99US-0138847;	PR	11-AUG-1999;	99US-0148319;
PR	14-JUN-1999;	99US-0139119;	PR	12-AUG-1999;	99US-0148341;
PR	16-JUN-1999;	99US-0139452;	PR	13-AUG-1999;	99US-0148565;
PR	16-JUN-1999;	99US-0139453;	PR	13-AUG-1999;	99US-0148684;
PR	17-JUN-1999;	99US-0139492;	PR	16-AUG-1999;	99US-0149368;
PR	17-JUN-1999;	99US-0139454;	PR	17-AUG-1999;	99US-0149175;
PR	18-JUN-1999;	99US-0139455;	PR	18-AUG-1999;	99US-0149426;
PR	18-JUN-1999;	99US-0139456;	PR	20-AUG-1999;	99US-0149722;
PR	18-JUN-1999;	99US-0139457;	PR	20-AUG-1999;	99US-0149722;
PR	18-JUN-1999;	99US-0139458;	PR	20-AUG-1999;	99US-0149722;
PR	18-JUN-1999;	99US-0139459;	PR	20-AUG-1999;	99US-0149722;
PR	18-JUN-1999;	99US-0139460;	PR	20-AUG-1999;	99US-0149722;
PR	18-JUN-1999;	99US-0139461;	PR	20-AUG-1999;	99US-0149722;
PR	18-JUN-1999;	99US-0139462;	PR	20-AUG-1999;	99US-0149722;
PR	18-JUN-1999;	99US-0139463;	PR	23-AUG-1999;	99US-0149902;
PR	18-JUN-1999;	99US-0139750;	PR	23-AUG-1999;	99US-0150566;
PR	18-JUN-1999;	99US-0139763;	PR	25-AUG-1999;	99US-0150566;
PR	21-JUN-1999;	99US-0139817;	PR	26-AUG-1999;	99US-0150884;
PR	22-JUN-1999;	99US-0139889;	PR	27-AUG-1999;	99US-0151065;
PR	23-JUN-1999;	99US-0140353;	PR	27-AUG-1999;	99US-0151086;
PR	23-JUN-1999;	99US-0140354;	PR	27-AUG-1999;	99US-0151086;
PR	24-JUN-1999;	99US-0140695;	PR	30-AUG-1999;	99US-0151303;
PR	24-JUN-1999;	99US-0140823;	PR	31-AUG-1999;	99US-0151438;
PR	29-JUN-1999;	99US-0140991;	PR	01-SEP-1999;	99US-0151930;
PR	30-JUN-1999;	99US-014187;	PR	01-SEP-1999;	99US-0152363;
PR	01-JUL-1999;	99US-0141842;	PR	07-SEP-1999;	99US-0153070;
PR	01-JUL-1999;	99US-0142154;	PR	10-SEP-1999;	99US-0153758;
PR	02-JUL-1999;	99US-0142055;	PR	13-SEP-1999;	99US-0154018;
PR	06-JUL-1999;	99US-0142390;	PR	15-SEP-1999;	99US-0154018;
PR	08-JUL-1999;	99US-0142820;	PR	16-SEP-1999;	99US-0154779;
PR	09-JUL-1999;	99US-0142920;	PR	20-SEP-1999;	99US-015

[illegible]

PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130049.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134766.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 28-MAY-1999; 99US-0136392.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142290.
 PR 09-JUL-1999; 99US-0142803.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147260.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149375.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151103.
 PR 31-AUG-1999; 99US-0151138.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0155659.
 PR 29-SEP-1999; 99US-0155659.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0158025.
 PR 07-OCT-1999; 99US-0158025.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 56.3%; Score 994; DB 21; Length 189;
Best Local Similarity 99.5%; Pred. No. 2, 8e-77;

Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 148 MSSTFPPLRNGIPSSLPRLSRISNCPYTPVSSPTSKPKPLPMWESIAKOSMAIAKQ 207
1 MSSTFPPLRNGIPSSLPRLSRISNCPYTPVSSPTSKPKPLPMWESIAKOSMAIAKQ 60
DB 1
OY 208 SNASFPYPPYAVASAPASPTHRHOFHTLATITPECDSDSSVDSGHWISQKFAQQOPFSA 267
61 SNASFPYPPYAVASAPASPTHRHOFHTPATITPECDSDSSVDSGHWISQKFAQQOPFSA 120
OY 268 SNVPTSPPTNLVVPARQKSPNTPAARFOEIGQSEPKPENSOVPRWEGERTHDVGMEDLEL 327
121 SNVPTSPPTNLVVPARQKSPNTPAARFOEIGQSEPKPENSOVPRWEGERTHDVGMEDLEL 180
DB 121
OY 328 TLGNKRGARG 336
181 TLGNKRGARG 189
DB 181
RESULT 9
AAG30139
ID AAG30139 standard; Protein: 284 AA.
XX AAG30139;
AC AAG30139;
XX 17-OCT-2000 (first entry)
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 35979.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35979.
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX RP1033405-A2.
PN RP1033405-A2.
XX 06-SEP-2000.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128719.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138547.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139859.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999;

```
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149358.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151338.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
```

```
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 30.7%; Score 541; DB 21; Length 284;
Best Local Similarity 40.2%; Pred. No. 3,8e+38;
Matches 143; Conservative 29; Mismatches 64; Indels 120; Gaps 17;

```
QY 23 RKSSWRERENRRERRRAVAIAKIYTGTPAQGDYNI PKHCDDNNEVUKALCEVAGWVBE 82
DB 6 RPTWKERENKRRERRRAIAKIIPAGLRHGFKL PKHCDDNNEVUKALCEVAGWVBE 65
QY 83 DGTYYRKCKPLPELAIAGTSRYTPYS-SONQSPLSAFOSPISYQVSPSSSEFPSPSR 141
DB 66 DGTYYRKCKPLPELAIAGTSRYTPYS-SONQSPLSAFOSPISYQVSPSSSEFPSPSR 117
QY 142 --GEPNNMSSTFFPLRN--GGIPSLPLSR--ISNCPVTPPVSSPTSKPKPLPNWES 196
DB 118 PFGDAN-----SLIPWLKLNSSNSPSTLPFFHGNSISAPVTPPLA----- 157
QY 197 IAKQSMALIAKQSMASFYFPAVASAPASPTRRHQFHLATIPCEDESDSTVDGHWISF 256
DB 158 -----RSPF-RDQ-----VITP-----DSG-WLS- 174
QY 257 QKFAQQQPSASWPTSPFTENLV-----KPA-POOM 286
DB 175 ---GMQTPQSG---PSSPFSLSVRNPFPDKEAFKMDSCNPMKTPQSGNCSFAIRPGV 228
QY 287 SPYTAAPORIGQSSSEPKF-----ENSQVKEWGERIH-DVGHEDELTTGNGRAR 335
DB 229 DQNSDVPMDGWTAEAFGCMNAAMANGWYFMEGERIHGCVSDDELTTIGNSRTR 284
```

RESULT 10
AAGS2728
ID AAGS2728 strand; Protein; 284 AA.
XX AAGS2728;
AC AAGS2728;
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 67061.
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160983.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 30.7%; Score 541; DB 21; Length 284;
Best Local Similarity 40.2%; Pred. No. 3,8e-38;

Matches 143; Conservative 29; Mismatches 64; Indels 120; Gaps 17;

QY 23 RKPSMRERNNRRRRRAVAAKIYGLRAQGDYMLPKDCNNNEVLKALCNEAGWYVE 82
DB 6 KTTWKEKRENNKRRRRRAIAAKIFAGLRHGNFKLPKDCNNNEVLKALCNEAGWYVE 65
QY 83 DGTYYKCKPLDPELIGTSSRVTPYS-SQNSPFLSASFQSPISQVSPSSSFPSPSR 141
DB 66 DGTYYKCKCPM-----DRMDLWNGSTISASFCSSYQSPRASYNPSSSSFPSPPTN 117
QY 142 --GEPNNNSSTFFPLRN--GGIPSLPSLR--ISNSCPYTPPVSPSTSKNPKPLPMWES 196
DB 118 PFGDAN-----SLIPWLKNLSNSPSKLPFHNGISIAPTTPPA----- 157
QY 197 IAKQSAIAIKQSAIAFYPPYAVSAPASPTHRHQHTLATIPCEDESSTVDSGHWISF 256
DB 158 -----RSPR-RDQ-----VTIP-----DSG-WLS- 174
QY 257 OKFAQOCPFSASVWPTSPFNLY-----KPA-PQOM 286
DB 175 ---GKGTPOSG---PSSPTSLVSKNPFDPKXAFKKGDCNSPMWTPQSGNCSPAIPAGV 228
QY 287 SPYTAAPQEIQSSSEKPF-----ENSOVQPMWGERIH-DVGMEDLELTGNGKAR 335
DB 229 DQNSDVPMADGMATAEAFGCMNAAMANGVPMWGERIHGECVSDDELITGNSRTR 284

RESULT 11
ID AAG29487
AC AAG29487
XX AAG29487;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment 890 ID NO: 35094.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; Genetic mapping; gene expression control; promoter;
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132843.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0143297.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144335.	PR	13-OCT-1999;	99US-0159295.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159684.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160776.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160778.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161920.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161922.
PR	04-AUG-1999;	99US-0147302.	PR	28-OCT-1999;	99US-0161932.
PR	05-AUG-1999;	99US-0147260.	PR	29-OCT-1999;	99US-0162142.
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			

QY	10	SAAAAAAAAAARRRRRRRRRRRRRRRAVAKIYTGIRACGDYILPRGCDNEVL	69
DB	3	AGGGGGGGSSGRIPTWKERENNNKKRRRRRAIRAKIYSGIRACGNYLPRGCDNEVL	62
QY	70	KALCIEAGWVYBEDGTYRKCK--PUPGRIAGTSSVTPYSSONOSPLSAPFOSPIPS	126
DB	63	KALCIEAGWIVEDDGTTRKGFSGHQHFOELLOTSAOI--HOSNOVHNHOLFQ--VLH	117
QY	127	YQVSPSSSSFPBSPKCEPDNNNNSTFF	153
DB	118	LTTEVOSHPRBRVHLAMTETILHTFF	144

Query Match

Best Local Similarity: 18.7%; Score 310; DB 21; Length 169;

Matches 68; Conservative 23; Mismatches 48; Indels 8; Gaps 3;

RESULT 12
AAG30141
ID AAG30141 standard; Protein; 205 AA.
XX AAG30141;
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 35981.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137228.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137702.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139352.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145006.
PR 22-JUL-1999; 99US-0145008.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

```

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160777.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13 34; Score 235.5; DB 21; Length 205;
Best Local Similarity 32.14; Pred. No. 4,1e-12;
Matches 86; Conservative 18; Mismatches 53; Indels 111; Gaps 15;

```

```

DB 81 PT-RDQ-----VTIP-----DSG-WLS-----GMTPQSG---DSSPTFLSVSNPF 117
QY 280 -----KPA-PQWSPNTAFQEIQSGSEFKF-----ENSQ 308
DB 118 FDKEAFRMGDCNSPMWTPGSGCNSCPAI PACVDONSDVPWADGNTAFAPGCMMAANGM 177
QY 309 VKPMEGERIH-DVGMEDLELTNGKAR 335
DB 178 VKPMEGERIHGBCVSDLELTNGSRTR 205

RESULT 13
ID MAG52730 standard; Protein; 205 AA.
AC MAG52730;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67063.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123160.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0132856.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 25-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.

```

PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	18-JUN-1999;	99US-0139492.	PR	16-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	17-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139456.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139459.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139461.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139462.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139463.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	99US-0139899.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	99US-0140354.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140695.	PR	10-SEP-1999;	99US-0153070.
PR	28-JUN-1999;	99US-0140823.	PR	13-SEP-1999;	99US-0153758.
PR	29-JUN-1999;	99US-0140991.	PR	15-SEP-1999;	99US-0154018.
PR	30-JUN-1999;	99US-0141287.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0141842.	PR	20-SEP-1999;	99US-0154779.
PR	01-JUL-1999;	99US-0141854.	PR	22-SEP-1999;	99US-0155133.
PR	02-JUL-1999;	99US-0142055.	PR	23-SEP-1999;	99US-0155486.
PR	06-JUL-1999;	99US-0142390.	PR	24-SEP-1999;	99US-0155659.
PR	08-JUL-1999;	99US-0142803.	PR	28-SEP-1999;	99US-0156458.
PR	09-JUL-1999;	99US-0142920.	PR	29-SEP-1999;	99US-0156596.
PR	12-JUL-1999;	99US-0142977.	PR	04-OCT-1999;	99US-0157117.
PR	13-JUL-1999;	99US-0143342.	PR	05-OCT-1999;	99US-0157753.
PR	14-JUL-1999;	99US-0143624.	PR	06-OCT-1999;	99US-0157865.
PR	15-JUL-1999;	99US-0144005.	PR	07-OCT-1999;	99US-0158023.
PR	16-JUL-1999;	99US-0144085.	PR	08-OCT-1999;	99US-0158229.
PR	16-JUL-1999;	99US-0144332.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145102.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145291.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145291.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145291.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145291.	PR	25-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145551.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	03-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	04-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161920.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161920.
PR	05-AUG-1999;	99US-0147302.	PR	28-OCT-1999;	99US-0161920.
PR	05-AUG-1999;	99US-0147302.	PR	28-OCT-1999;	99US-0161920.
PR	06-AUG-1999;	99US-0147260.	PR	28-OCT-1999;	99US-0161920.
PR	06-AUG-1999;	99US-0147303.	PR	28-OCT-1999;	99US-0161920.
PR	06-AUG-1999;	99US-0147416.	PR	28-OCT-1999;	99US-0161920.
PR	09-AUG-1999;	99US-0147493.	PR	29-OCT-1999;	99US-0162142.

Query Match 113.3%; Score 235.5; DB 21; Length 205;
 Best Local Similarity 32.1%; Pred. No. 4.1e-12;
 Matches 86; Conservative 18; Mismatches 53; Indels 11; Gaps 15;
 QY 110 SONQSPLSAFOSPRISYQVSPSSSFPSPSR--GEINNMSTFPFLRN--GGIPSSL 165

[illegible]

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136682.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2003, 14:24:08 ; Search time 20 Seconds
(without alignments)
1615.634 Million cell updates/sec

Title: US-09-995-938a-7

Perfect score: 1765
Sequence: 1 MTSDDGATSTSAIAAAAAAAAAA.....IHDVGMEDLELTGNGKARG 336

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 76:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1726.5	97.8	333	2	G96780	unknown protein F9
2	1519.5	86.1	335	2	G86326	protein P18014.7 l
3	825	46.7	276	2	T46152	hypothetical prote
4	585.5	33.2	325	2	H96815	hypothetical prote
5	541	30.7	284	2	T05026	hypothetical prote
6	203.5	11.5	687	2	T02459	probable beta-amy1
7	146	8.3	1859	1	A34092	DNA-directed RNA p
8	146	8.3	1859	1	T29959	DNA-directed RNA p
9	137.5	7.8	383	2	T46707	proteophosphoglyc
10	137.5	7.8	528	2	T47141	Gastric mucin (clo
11	132	7.5	493	2	S50755	hypothetical prote
12	132	7.5	493	2	S46650	hypothetical prote
13	131	7.4	467	2	A27677	weta protein - pen
14	131	7.4	1932	1	A28490	DNA-directed RNA p
15	131	7.4	1970	1	S21054	DNA-directed RNA p
16	131	7.4	1970	1	T38186	RNA polymerase II
17	128	7.3	317	2	S55316	mucin (clone PGM-2
18	127.5	7.2	514	2	T39903	serine-rich protei
19	127.5	7.2	1493	2	T110757	MAP kinase kinase
20	126.5	7.2	1733	1	RNBY2L	DNA-directed RNA p
21	125.5	7.1	333	1	JM0097	bicoid-related hom
22	125	7.1	2232	2	T43434	hypothetical prote
23	124.5	7.1	1896	1	RNPF2L	hypothetical prote
24	124	7.0	1367	1	S48478	glucan 1,4-alpha-g
25	123.5	7.0	268	2	H84684	En/Spm-like transp
26	123	7.0	989	2	T02568	hypothetical prote
27	122.5	6.9	977	2	S14183	DNA-directed RNA p
28	122.5	6.9	2414	2	A54277	transcription adap
29	122	6.9	322	2	S25299	extensin precursor

30	122	6.9	428	2	T43532	zinc finger protei
31	122	6.9	1491	2	S14182	DNA-directed RNA p
32	122	6.9	532	2	S74453	hypothetical prote
33	121.5	6.9	855	2	T41336	probable nitroge
34	121	6.9	1351	2	S20754	hypothetical prote
35	121	6.9	1752	1	S26849	DNA-directed RNA p
36	121	6.9	1834	1	JDMU1	DNA-directed RNA p
37	121	6.9	1840	2	G85422	hypothetical prote
38	120.5	6.8	902	2	A56823	DNA-directed RNA p
39	119.5	6.8	650	2	S14181	DNA-directed RNA p
40	119	6.7	625	2	T07796	DNA-directed RNA p
41	118.5	6.7	1032	2	T34433	hypothetical prote
42	118	6.7	596	2	T03908	hypothetical prote
43	118	6.7	2090	2	S26058	probable transform
44	118	6.7	2845	2	I49505	adenomatous polyo
45	117	6.6	1365	2	S14871	suppressor two of

ALIGNMENTS

RESULT 1		
G96780		
unknown protein F9E10.7 (imported) - Arabidopsis thaliana		
C/Species: Arabidopsis thaliana (mouse-ear cress)		
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001		
C/Accession: G96780		
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Gonn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Ansen, N.F.; Hughes, B.; Hultzar, L.		
Nature 408, 816-820, 2000		
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lutero, J.S.; Maiti, R.; Marshall, R.; Rooney, T.; Rowley, D.; Sakano, H.		
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.		
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.		
A/Reference number: A86141; MUID:21016719; PMID:11130712		
A/Accession: G96780		
A/Status: preliminary		
A/Molecule type: DNA		
A/Residues: 1-333 <STO>		
A/Cross-references: GB:AE005173; NID:96646757; PIDN:AAF21069.1; GSPDB:GN00141		
C/Genes: F9E10.7		
A/Map position: 1		
Query Match		97.8%; Score 1726.5; DB 2; Length 333;
Best Local Similarity		98.8%; Pred. No. 5e-110;
Matches 333; Conservative 0; Mismatches 1; Indels 3; Gaps 1;		
QY	1	MTSDGATSTSAIAAAAAAAAAARRRPSRRERENRRRRRAVAKIYTGIRAGDYNLP 60
DB	1	MTSDGATSTSAIAAAAAAAAAARRRPSRRERENRRRRRAVAKIYTGIRAGDYNLP 60
QY	61	KKCDNNEVKALCVBAWVVEEDGTYRKCKPLPGSIAGTSSVTVYSSQNOPLSSAF 120
DB	61	KKCDNNEVKALCVBAWVVEEDGTYR---KPLPGSIAGTSSVTVYSSQNOPLSSAF 117
QY	121	QSPFIPSYOVSPSSSPSPSPSGEPNNMMSGTFPPFLNNGIIPSLPLRLISNCPVMPVPV 180
DB	118	QSPFIPSYOVSPSSSPSPSPSGEPNNMMSGTFPPFLNNGIIPSLPLRLISNCPVMPVPV 177
QY	181	SSPTSKPKPLPMBESIAKQSMASFNYPYAVASAPTRHGFHTLATIPEC 240
DB	178	SSPTSKPKPLPMBESIAKQSMASFNYPYAVASAPTRHGFHTLATIPEC 237
QY	241	DESDDSTVSGHWISFOKFAQQOFPASWVPTSPTRILVVPAAQOQSPNTAAFOEICQSS 300
DB	238	DESDDSTVSGHWISFOKFAQQOFPASWVPTSPTRILVVPAAQOQSPNTAAFOEICQSS 297
QY	301	EFKFNQVQKPEGERIHDVGMEDLELTGNGKARG 336
DB		

DB 298 EFKFNSQVQKPEGRHIDVGMEDLETLGNGKARG 333

RESULT 2

G86326

protein P18014.7 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear crease)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: G86326

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroe, J.S.; Malt, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

ARTICLE: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; PMID:11130712

A/Accession: G86326

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-335 <STO>

A/Cross-references: GB:A8005172; NID:g8778414; P1DN:AAF79422.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

A/Map position: 1

R/Querier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salenou

submitted to the Protein Sequence Database, May 1999

A/Reference number: Z16409

A/Accession: T08396

A/Molecule type: DNA

A/Residues: 1-276 <QUB>

A/Cross-references: EMBL:AL049862; GSPDB:GN00061; ATSP:F18B3.30

A/Experimental source: cultivar Columbia; BAC clone F18B3

C/Genetics:

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

A/Map position: 3

us-09-995-938a-7.rpr

```

QY 23 RKPMSRENNRRRRRAVAAKIYTGIRAGDYNLPKHCDDNNEVLKALCVAEGWVEE 82
DB 6 RMPYMERENNRKRERRRAIAKIPTGLRMVGYELPKHCDDNNEVLKALCVAEGWVEE 65
QY 83 DGTTRKGC-KPLPG-ELIGTSRVTPTYSQNSPLSAFOSPIPSYQVSPSSSPSPS 140
DB 66 DGTTRKGCSPRVERMEIGCGSATPSCSSYQSPCASVYPSGNSFMSPASSPANTL 125
QY 141 RCEPNNMSTFPPLFRNGIPSS---LQSLRSN-----SCPVTTPVSSPTSKIP 188
DB 126 SGD-----GOSLIPWIKHLSTSSASSSSSRUPNLYIPGGSISAPVYPLPSSPTARTP 180
QY 189 KPLPWMSIAKQSMALAKQSMASFNYPFAVASAPATHRHOPHTLATIPCEDESSTV 248
DB 181 RKNTDWQOL-----NNSFVSSSTPSPPTQ-----IIPDSE----- 211
QY 249 DSGHITSFOKPAQOQPFSAWPTSPFTPLVYKPAQOMSPTAAFOE----- 295
DB 212 ---WPSGIQLAQO-----VPASPTFLV-----SQNPFGFKEBAASAAGGGGSR 253
QY 296 ---IGQSS-----EFKE---ENSQVYPMGEGRH-DVYME 323
DB 254 MWTPOSGTCSFPAIPGADQTDADVPMSEAVALPEFAFGSNTGLVAMEGERIHESGSD 313
QY 324 DLELTGNGKAR 335
DB 314 DLELTGNSSTR 325

```

RESULT 5

```

105026
hypothetical protein F13C5_60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C/Accession: T05026
R:Bevan, M.; Pohl, T.; Weizenecker, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuel
submitted to the Protein Sequence Database, February 1998
A/Reference number: Z15395
A/Accession: T05026
A/Molecule type: DNA
A/Residues: 1-284 <BEV>
A/Cross-references: EMBL:AL021711
A/Experimental source: cultivar Columbia; BAC clone F13C5
C/Genetics:
A/Map position: 4
A/Intons: 72/3
A/Note: F13C5_60

```

```

Query Match 30.7% Score 541; DB 2; Length 284;
Best Local Similarity 40.2% Pred. No. 1.1e-29;
Matches 143; Conservative 29; Mismatches 64; Indels 120; Gaps 17;
QY 23 RKPMSRENNRRRRRAVAAKIYTGIRAGDYNLPKHCDDNNEVLKALCVAEGWVEE 82
DB 6 RMPYMERENNRKRERRRAIAKIPTGLRMVGYELPKHCDDNNEVLKALCVAEGWVEE 65
QY 83 DGTTRKGCPLPGEIAGTSRVTPTYS-QNQSPLSAFOSPIPSYQVSPSSSPSPSR 141
DB 66 DGTTRKGCSPM-----DRMDLMNGSTASPCSSVQHPSPRASYNSSPSSSPSTN 117
QY 142 --GEPNNMSTFPPLRN-GGIPSLPSLR-ISNCPVTTPVSSPTSKIPKPLWMS 196
DB 118 PFGGAN-----SLIPWLKNSNSPSKLPFRHNSISAPVTPPLA----- 157
QY 197 IAKQSMALAKQSMASFNYPFAVASAPATHRHOPHTLATIPCEDESSTVDSGWISF 256
DB 158 -----RSPT-RDQ-----VILP-----DSG-WLS- 174
QY 257 OKFAQOQPFSAWPTSPFTPLV-----KPA-PQM 286
DB 175 ---GMQTPQSG---PSSPTFLVGRNPPFDKAFPMQDCNSPMWTPQSGNCSFAIRAGV 228
QY 287 SPNTAAFOEIGQSSSEPKF-----ENQVYPMGEGRH-DVYMEDLELTGNGKAR 335

```

```

DB 229 DQNSDVPMADGWTAEPAFGCNMAAAGWYKPMGEGRHGEVSDDELTGNSRTR 284

```

RESULT 6

```

102459
probable beta-amylase At2g45880 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 02-Feb-2001
C/Accession: T02459; A04896
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kau
submitted to the EMBL Data Library, August 1998
A/Description: Arabidopsis thaliana chromosome II BAC F4118 genomic sequence.
A/Reference number: Z14674
A/Accession: T02459
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-687 <ROU>
A/Cross-references: EMBL:AC004665; NID:G3386593; PID:G3386606
A/Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Genito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Xoo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A04420 MUID:20083487; PMID:10617197
A/Accession: A04896
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-687 <STO>
A/Cross-references: GB:AB002093; NID:G3386606; PID:AA028536.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g45880; F4118.14
A/Map position: 2
A/Intons: 138/3; 229/3; 255/3; 323/3; 388/3; 443/3; 530/3; 603/3

```

```

Query Match 11.5% Score 203.5; DB 2; Length 687;
Best Local Similarity 39.5% Pred. No. 2.3e-06;
Matches 51; Conservative 16; Mismatches 53; Indels 9; Gaps 3;

```

```

QY 21 ARKPSRENNRRRRRAVAAKIYTGIRAGDYNLPKHCDDNNEVLKALCVAEGWV 80
DB 67 SRSRPPEKERTKLRERHRAITARIIGLRPHGNVYLVRADINDVIALAREXGVV 126
QY 81 EEDGTTY---RKQCKPLGEIA-GTSSRVTPTYSQNSP-----LSSAFOSPIPSYQVSP 131
DB 127 LPDGTTFPSKSGRTKPTGSSAVALAGSSASHIASOKTSPALRVVSSGLRSEVELSSCRM 186
QY 132 SSSSPSPS 140
DB 187 KGVFTPAVS 195

```

RESULT 7

```

A34092
DNA-directed RNA polymerase (EC 2.7.7.6) II large chain - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: A34092
R:Bird, D.M.; Riddle, D.L.
Mol. Cell. Biol. 9, 4119-4130, 1989
A/Title: Molecular cloning and sequencing of ama-1, the gene encoding the largest subunit
A/Reference number: A34092; MUID:90066416; PMID:2566513
A/Accession: A34092
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1859 <BIR>
A/Cross-references: GB:M29295; NID:G156404; PID:AAA28126.1; PID:G156405
A/Note: the authors failed to translated the codon GAG for residue 917 as Glu, and CAT E
C/Superfamily: human DNA-directed RNA polymerase II largest chain
C/Keywords: DNA binding; nucleocytoplasmic transferase; tandem repeat; transcription; zinc finger
Query Match 8.3% Score 146; DB 1; Length 1859;

```



```

RESULT 11
S50755
hypothetical protein VSP-3 - Chlamydomonas reinhardtii
C/Species: Chlamydomonas reinhardtii
C/Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 21-Jul-2000
C/Accession: S50755
R/Reviewer: J.P.; Molendijk, A.J.; van Egmond, P.; Klis, F.M.; Goodenough, U.W.; Haring
Plant Mol. Biol. 26, 947-960, 1994
A/Title: Domain conservation in several volvocalean cell wall proteins.
A/Reference number: S50754; MUID:95093034; PMID:8000007
A/Accession: S50755
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-473 <MOB>
A/Cross-references: EMBL:129029; NID:G530875; PIDN:AA853953.1; PID:G530876

```

RESULT 12
S46660
weta protein - *Penicillium chrysogenum*
C/Species: *Penicillium chrysogenum*
C/Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 26-Aug-1999
C/Accession: S46660
R/Prade, R.A.; Timberlake, W.E.
Mol. Gen. Genet. 244, 539-547, 1994
A/Title: The *Penicillium chrysogenum* and *Aspergillus nidulans* weta developmental regula
A/Reference number: S46660; MUID:94359480; PMID:8078481
A/Accession: S46660
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-499 <PRA>
A/Cross-references: EMBL:X80058; NID:9510873; PIDN:CAAS6364.1; PID:9510874
C/Superfamily: regulatory protein weta

Oy		42	--AVAKIYTGLRAGGDVYLPRKHGDNHVLKALCVENGWVEEDGTTYRKCKPLPGEI-	98
Dd		166	NLSKRMAYPEMAQRLQNFSLH-----GSEDRLPLSPPPSDVL	204
Oy		99	-----AGTSRVTPTYSSQ--NQSLPSAFOSPIPS-----YGVSP	131
Dd		205	IQHENVPTREQINHQHGDSASERASQYDARLYQQSP---SVMSPESTIMASAROQGHYTQP	261
Oy		132	SSSSPFSSSRGRGEPNNNMSTFFPFLNGCIPSSLFELRI---SNSCPVTTPVASSPTSK-	186
Dd		262	SSSSLTJNSSPSSSADDIFSSSS-----HSSDPHSLSSWQSOPDLHAASSLSFTFDLGGQDSQW	315
Oy		187	-----NPKPLPNMESIAKQSAIAIKQSNASHNYPRYAASA-----	221
Dd		316	WSMPESRVAAQQAQAVLTSEPTPVRTMOSVGSDNDMMQGGLTIQFN-PDYDMSADHSFSSSN	374
Oy		222	--PASPT-----HRHOEHTLATIPECDD-ESDSSTVDSCGWISFQKFAAQOOPFGASWVPT	272
Dd		375	MLPATPQKFDTSGFNISQVHNVRSPSLSPKAGTSPBDTRKMIIHFQ--AHSSPHSQRL-S	431
Oy		273	SPTPLULVPAPAOQNSPNTAIFOEIQGQSEFFKEKNQSVKKWEGERIHQDVGMEDLELTIGNG	332
Dd		432	GQSMNAAPKPAPABGSSRGSRGNKSVS---EVNFYA-----HDSKILTELPGSS	479
Oy		333	KAR	335
Dd		480	KTK	482

RESULT 13
A27677
DNA-directed RNA polymerase (EC 2.7.7.6) II largeat chain - Chinese hamster (Fragment)
-Sequences Cytocool

A;Residues: 1-467 <ALL>
A;Cross-references: GB:M9538; NID:G191170; PID:AA37006.1; PID:G387055
C;Genetics:
A;Gene: RPO21
A;Introns: 33/1; 83/1
C;Superfamily: human DNA-directed RNA polymerase γ 1 largest chain
C;Keywords: DNA binding; nucleic acid; transcriptase; tandem repeat; transcription; zinc fi

RESULT 14
A28490
DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain - mouse

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2003, 14:24:08 ; Search time 17 Seconds

(without alignments)
929.469 Million cell updates/sec

Title: US-09-995-938A-7

Perfect score: 1765
Sequence: 1 MTSNGATSTSTAAAAA.....IHDVGMEDLELTGNGKARG 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	146	8.3	1852	1	RPB1_CAEEL
2	133.5	7.6	314	1	PIX2_BRAR
3	132	7.5	499	1	WETA_PENCH
4	131	7.4	467	1	RPB1_CRGR
5	131	7.4	1970	1	RPB1_HUMAN
6	131	7.4	1970	1	RPB1_MOUSE
7	128.5	7.3	326	1	PIX2_XENLA
8	127.5	7.2	1493	1	MKX1_RAT
9	126.5	7.2	317	1	PIX2_HUMAN
10	126.5	7.2	317	1	PIX2_MOUSE
11	126.5	7.2	324	1	PIX2_RAT
12	126.5	7.2	1733	1	RPB1_YEAST
13	125.5	7.1	333	1	PIX2_CHICK
14	124.5	7.1	1887	1	RPB1_DROME
15	124	7.0	1367	1	AMYH_YEAST
16	123.5	7.0	917	1	SMOQ_HUMAN
17	122.5	6.9	2414	1	P300_HUMAN
18	121.5	6.9	855	1	GAF1_SCHPO
19	121	6.9	806	1	MK07_MOUSE
20	121	6.9	1752	1	RPB1_SCHPO
21	121	6.9	1840	1	RPB1_ARATH
22	120.5	6.8	428	1	RSV1_SCHPO
23	120.5	6.8	902	1	MK1_DICDI
24	120.5	6.8	1493	1	MK1_MOUSE
25	120	6.7	1192	1	RTN4_HUMAN
26	118.5	6.7	845	1	KRUH_DROME
27	118	6.7	2090	1	N214_HUMAN
28	118	6.7	2842	1	APC_RAT
29	118	6.7	2845	1	APC_MOUSE
30	117.5	6.7	292	1	PIX3_XENLA
31	117	6.6	1365	1	SUZ2_DROME
32	117	6.6	5038	1	PCLO_MOUSE
33	116.5	6.6	565	1	SCR1_SCHPO

34	116	6.6	533	1	EGR1_MOUSE
35	116	6.6	531	1	YGF1_YEAST
36	115.5	6.5	898	1	YGB4_YEAST
37	115.5	6.5	1794	1	YDC9_SCHPO
38	115	6.5	697	1	AMS2_SCHPO
39	115	6.5	815	1	MK07_HUMAN
40	114.5	6.5	1596	1	GLI3_MOUSE
41	114.5	6.5	1874	1	POLR_KRWY
42	113	6.4	817	1	VRP1_YEAST
43	112.5	6.4	720	1	GTL1_SCHPO
44	112.5	6.4	1495	1	MKX1_HUMAN
45	111.5	6.3	413	1	ESL1_SCHPO

ALIGNMENTS

RESULT 1
ID RPB1_CAEEL STANDARD; PRT; 1852 AA.
AC P16356; Q20090;
DT 01-AUG-1990 (Rel. 15, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase II large subunit (EC 2.7.7.6).
GN AWA-1 OR F3644.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=90066416; PubMed=2586513;
RA Bird D.M., Riddle D.L.;
RT "Molecular cloning and sequencing of awa-1, the gene encoding the
RT largest subunit of Caenorhabditis elegans RNA polymerase II,";
RL Mol. Cell. Biol. 9:4119-4130(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Miller N., Bradshaw H.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (JUG-2001) to the EMBL/GenBank/DBJ databases.
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N-nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM NINE TO
CC FOURTEEN DIFFERENT POLYPEPTIDES.
CC -1- SUBCELLULAR LOCATION: Nucleus.
CC -1- PPM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC THE PHOSPHORYLATION ACTIVATES POL2.
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND TRNA GENES.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M29235; AAA28126.1; -;
DR EMBL; U53333; AAA96158.2; -;

	PIR	A34092; A34092.	
DR	Momped:	F36A4.7; CE28300.	
DR	InterPro:	IPIR000722; RNA_pol_A.	
DR	InterPro:	IPIR007080; RNA_pol_Rpb1_1.	
DR	InterPro:	IPIR007066; RNA_pol_Rpb1_3.	
DR	InterPro:	IPIR007083; RNA_pol_Rpb1_4.	
DR	InterPro:	IPIR007081; RNA_pol_Rpb1_5.	
DR	InterPro:	IPIR007075; RNA_pol_Rpb1_6.	
DR	InterPro:	IPIR007073; RNA_pol_Rpb1_7.	
DR	InterPro:	IPIR006592; RNA_pol_A.	
DR	InterPro:	IPIR00684; RNA_polIT_repeat.	
DR	pfam:	PF04997; RNA_pol_Rpb1_1.	
DR	pfam:	PF00623; RNA_pol_Rpb1_2; 1.	
DR	pfam:	PF04983; RNA_pol_Rpb1_3; 1.	
DR	pfam:	PF05000; RNA_pol_Rpb1_4; 1.	
DR	pfam:	PF04998; RNA_pol_Rpb1_5; 1.	
DR	pfam:	PF04992; RNA_pol_Rpb1_6; 1.	
DR	pfam:	PF05001; RNA_pol_Rpb1_7; 1.	
DR	SMART:	SMD0663; RPOLA_N. 1.	
DR	PROSITE,	PS00115; RNA_POL_II_REPEAT; 26.	
KW	Transferase;	DNA-directed RNA polymerase; Transcription; zinc; Repeat;	
KM	DNA-binding;	Nuclear protein; Phosphorylation; Zinc-finger.	
FT	ZN FING	66 82 C2H2-TYPE (POTENTIAL).	
FT	DOMAIN	1560 1852 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.	
FT	CONFILCT	215 215 V -> D (IN REF. 1).	
FT	CONFILCT	911 911 R -> RVAVQAQATLC (IN REF. 1).	
FT	CONFILCT	959 959 I -> D (IN REF. 1).	
FT	CONFILCT	974 974 Q -> L (IN REF. 1).	
FT	CONFILCT	990 991 KP -> NA (IN REF. 1).	
FT	CONFILCT	1156 1158 MISSING (IN REF. 1).	
FT	CONFILCT	1402 1403 IT -> YG (IN REF. 1).	
SO	SEQUENCE	1852 AA; 203978 MW; 211E4E563119088B CRC64;	
<hr/>			
Query Match 8.3%; Score 146; DB 1; Length 1852;			
Best Local Similarity 27.5%; Pred. No. 0.055;			
Matches 67; Conservative 34; Mismatches 95; Indels 48; Gaps 12;			
OY	85	TTRYKGCPLPEELIGTSSRVTPYSQNQSPLSAFOSPIRPYOYSPSSSSSTP-----PPPS 140	
DB	1539	ISPAAGCPVPBPSPBSPTSPSVSP-TSPSYSPSPSPSPSPSY--SPSPSPSPSPSPSY 1655	
OY	141	RGEPNNNNSSTFPFLRNCGISSLPSLR-RINSQCV-----TPVSSPTSKNPK 189	
DB	1656	PSBPSPSPSPS-----PSYGSPSRFYSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 1704	
OY	190	PL-PMMESIAKOSMAIAKOSMASFNV---PFAVASAPA-SPTHRHOFHTLATIPEDCES 243	
DB	1705	PRSPESYESGGCSP 1766	
OY	244	DGSIVDSGHMWISFOKRAQCQPSPASKNVPTSPFFNLVKAPQQMSPTTAAF-----QEIQ 298	
DB	1765	VNPPTSRCGF-----SSPYSPSPSPSPSPSPSPSYT--PSPDYSPSPSPSPSPSPSEDOGT 1816	
OY	239	SSEP 302 ::: ::: :::	
DB	1817	SNQY 1820	
<hr/>			
RESULT 2			
ID	PIX2_BRARE	STANDARD; PRT; 314_AA.	
AC	O9WSZ2; O9WSZ1;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Pituitary homeobox 2. PIRX2 OR PIIX2A.		
OS	Brachydanio rerio (zebrafish) (Danio rerio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
OX	NCBI_Taxid=7955;		

RP	[1]	SEQUENCE FROM N.A. (ISOFORMS PITX2A AND PITX2C).	
RA	Yan Y.-T., Gitman K., Ding J., Burdine R.D., Corrales J.,		
RA	Price S.M., Talbot W.S., Schier A.F., Shen M.M.,		
RT	"Conserved role for EGF-CFC genes in vertebrate left-right axis		
RT	formation.";		
RL	Submitted (JUN-1999) to the EMBL/GenBank/DDJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORMS PITX2A AND PITX2C).		
RA	Eisner J.J., Zhang J., Yost H.J.,		
RT	"Conserved isoforms of the pitx2 gene have differential activity and		
RT	expression in zebrafish.";		
RL	Submitted (MAR-1999) to the EMBL/GenBank/DDJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A. (ISOFORM PITX2A).		
RX	MEDLINE=99432240; PubMed=10500184;		
RA	Takui T., Capdevila J., Tamura K., Ruiz-Lozano P.,		
RA	Rodriguez-Esteban C., Yonei-Tamura S., Magallon J.,		
RA	Chandrasekara R.A.S., Chien K., Blumberg B., Evans R.M.,		
RA	Izpisua-Belmonte J.-C.,		
RT	"Multiple left-right asymmetry defects in Shh(-/-) mutant mice unveil		
RT	a convergence of the Shh and retinoic acid pathways in the control of		
RT	Lefty-1.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 96:11376-11381(1999).		
CC	-1- FUNCTION: May play an important role in development and		
CC	maintenance of anterior structures.		
CC	-1- SUBCELLULAR LOCATION: Nuclear.		
CC	-1- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;		
CC	Name=PITX2C;		
CC	Isoid=Q9W5Z2-1; Sequence=Displayed;		
CC	Name=PITX2A;		
CC	Isoid=Q9W5Z2-2; Sequence=VSP 002268, VSP 002269;		
CC	-1- SIMILARITY: BELONGS TO THE PAIRED HOMEBOX FAMILY. BICOID		
CC	SUBFAMILY.		
CC	-1- SIMILARITY: Contains 1 homeobox domain.		
CC	-1- SIMILARITY: Contains 1 OAR domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AF156905; AAD40179.1; -		
DR	EMBL; AF156906; AAD40180.1; -		
DR	EMBL; AF132446; AAD34390.1; -		
DR	EMBL; AF132447; AAD34391.1; -		
DR	EMBL; AF181681; AAF00486.2; -		
DR	HSSP; P06601; IFUL.		
DR	ZFIN; ZDB-GENE-990714-27; pitx2a.		
DR	InterPro; IPR003654; Homeo OAR.		
DR	InterPro; IPR003356; Homeobox.		
DR	InterPro; IPR007104; Paired_homeo.		
DR	PIfam; PF00046; homeobox; 1.		
DR	PIfam; PF03826; OAR; 1.		
DR	PRINTS; PR00024; HOMEBOX.		
DR	Prodom; PD00010; Homeobox; 1.		
DR	SMART; SM00389; HOX; 1.		
DR	PROSITE; PS00027; HOMEBOX_1; 1.		
DR	PROSITE; PS50071; HOMEBOX_2; 1.		
DR	PROSITE; PS50073; OAR; 1.		
KM	Homeobox; DNA-binding; Developmental protein; Nuclear protein;		
KM	Alternative splicing;		
FT	DNA BIND 82 141 HOMEBOX.		
FT	DOMAIN 276 289 OAR.		
FT	DOMAIN 282 286 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).		
FT	VARSPPLIC 1 45 MISSING (in isoform PITX2A).		
FT	FTID=VSP_002268.		
FT	FTID=VSPITSSPEV -> MSHCRKLASTCAOL (in		
FT	isoform PITX2A).		
FT	VARSPPLIC 46 60		

QY	2	TSDGATGTS	AAAAA	AAAAA	AAARPK	-----	SWREEMNR	RRRRRR	-----	41																						
Db	109	TAHGSLGTS	---	---	---	---	PTTPPATPRK	KPTQSALITPKS	IRHRS	PNERRSHLR	KOSFSP	LMRSS	165																			
QY	42	--	AVAKI	YTG	LRAG	DYNLPR	GHCDNNE	VLKAL	CVEAG	WVEEDG	TTYYR	GCKPLPGEI	98																			
Db	166	NLSKARMA	YPEMA	QRLN	FSLH	-----	-----	-----	-----	-----	-----	-----	204																			
QY	99	-----	-----	-----	AGTSR	TPP	ISSO	---	---	---	---	YQVSP	131																			
Db	205	IQHENN	PTQ	IMNO	HGD	SAERAS	QYDAR	LYOOS	P---	---	---	SVMS	PSPIAMS	AROOCHYIAQ	261																	
QY	132	SSSP	SPSSR	RGEP	PNNN	MSST	FFPLR	NGIP	SPSLRI	---	---	SNSCP	TPPV	VSSTK	186																	
Db	262	SSSL	STLNS	SPSS	ADIF	SSS	-----	---	---	---	---	HS	DPHLS	SSWOS	DP	LHASS	LSFT	P	LOG	DSOW	315											
QY	197	-----	-----	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	221									
Db	316	WSPMP	SRVA	QOQ	AYLT	PTPE	VRTQ	SGSD	NDMM	QGLM	QFN	---	---	---	---	---	---	---	---	---	---	---	374									
QY	222	--	PASPT	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	272									
Db	375	MLPAT	POK	FDT	SP	TSQ	OVHNV	SRSP	SLSP	PKAGT	SPRDR	KWIF	HO	---	---	---	---	---	---	---	---	431										
QY	273	SPTFLV	KAP	POQ	SPNT	APAF	OE	QGS	SE	FPF	ENS	QV	KP	W	SGER	I	HDV	G	ME	LE	L	TL	NG	332								
Db	432	GGSNN	APR	EP	AL	SSSS	SR	KSV	S	VSV	---	---	---	---	---	---	---	---	---	---	---	---	---	479								
QY	333	KAR	335	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----								
Db	480	KTK	482	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----								
RESULT 4																																
RPBI	CRIGR	STANDARD;	PRT,	467	AA.																											
AC	P11414;																															
DT	01-OCT-1989	(Rel. 12,	Created)																													
DT	01-OCT-1989	(Rel. 12,	Last sequence update)																													
DT	15-DEC-1998	(Rel. 37,	Last annotation update)																													
DS	DN	-directed	RNA polymerase II	largest	subunit	(EC 2.7.7.6)	(Rpb1)																									
DE	(Fragment).																															
GN	POLR2A.																															

CC between the Swiss Institute of Bioinformatics and the EMBL outcater -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).

CC -----
 DR EMBL; M1938; AAA37008.1; -
 DR PIR; A27677; A27677.
 DR InterPro: IPR000684; RNA polII repeat.
 DR Pfam: PF05001; RNA pol Rpb1 R; 28.
 DR PROSITE; PS00115; RNA_POL_II_REPEAT; 43.
 DR Transferrase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
 KW DNA-binding; Nuclear protein; Phosphorylation.
 FT NON TER 1
 FT DOMAIN 90 457 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
 SO SEQUENCE 467 AA; 48167 MW; 0AADFD92B5D2BCE CRC64;

Query Match 7.4%; Score 131; DB 1; Length 467;
 Best Local Similarity 29.7%; Pred. No. 0.11;
 Matches 65; Conservative 23; Mismatches 91; Indels 40; Gaps 12;

QY 93 PLPGELAGTSRRVPSYSSONQSPISAFQSPISYQVSPSSSSP-PSPSRGEPPNNMSST 151
 DB 261 PTSPSYSPSPNTP-TPSPNTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 314
 QY 152 FPFPLRNGCIPSSLSRLISNSCPTPVSSPTSKNKP-----LPNMESIA-KQSM 202
 DB 315 YTP-----SSPSYSPS---SPSYSPSPSKYPTSPSPSPSPSPSPSPSPSPSPSPSPSP 366
 QY 203 AIAKQSMASFN-----PPYAVASAP-SPTHHQHTLTATPECCSDSSSTYDS 250
 DB 367 TSPKYSPTSP 426
 QY 251 GHWISFQKFAQQCPFSAMVPTSPFNVKAPQOMSPN 289
 DB 427 -YSPTSPKSTYSP 461

RESULT 5

RPB1_HUMAN STANDARD; PRT; 1970 AA.

AC P24928;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6) (RPB1).
 GN POLR2A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92178992; PubMed=1542581;
 RA Wintzerich M., Acker J., Vicaire S., Vigneron M., Keding C.;
 RT "Complete sequence of the human RNA polymerase II largest subunit";
 RL Nucleic Acids Res. 20:910-910(1992).

CC SEQUENCE FROM N.A.
 CC MEDLINE=95347616; PubMed=7622068;
 CC Mita K., Teuji H., Morimyo M., Takahashi E., Nenoi M.,
 CC Ichimura S., Yamuchi M., Hongo B., Hayashi A.;
 CC "The human gene encoding the largest subunit of RNA polymerase II";
 CC Gene 159:285-286(1995).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -1- SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.

CC THE PHOSPHORYLATION ACTIVATES POL2.
 CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
 CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
 CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
 CC III FOR 5S AND TRNA GENES.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outcater -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).

CC -----
 DR EMBL; X63564; CA45125.1; -
 DR EMBL; X74874; CA52862.1; JOINED.
 DR EMBL; X74873; CA52862.1; JOINED.
 DR EMBL; X74872; CA52862.1; JOINED.
 DR EMBL; X74871; CA52862.1; JOINED.
 DR PIR; I38186; I38186.
 DR PIR; S21054; S21054.
 DR Gene; HGNC:9187; POLR2A.
 DR MIM; 180660; -

DR GO; GO:0005665; C:DNA-directed RNA polymerase II, core complex; TAS.
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003677; F:DNA binding activity; NAS.
 DR GO; GO:0003011; F:DNA-directed RNA polymerase II activity; NAS.
 DR GO; GO:0006355; F:regulation of transcription, DNA-dependent; NAS.
 DR GO; GO:0006366; P:transcription from pol II promoter; NAS.

DR InterPro: IPR000722; RNA_pol_A.
 DR InterPro: IPR007080; RNA_pol_Rpb1_1.
 DR InterPro: IPR007066; RNA_pol_Rpb1_3.
 DR InterPro: IPR007083; RNA_pol_Rpb1_4.
 DR InterPro: IPR007081; RNA_pol_Rpb1_5.
 DR InterPro: IPR007075; RNA_pol_Rpb1_6.
 DR InterPro: IPR007073; RNA_pol_Rpb1_7.
 DR InterPro: IPR006592; RNA_pol_A_N.
 DR InterPro: IPR00684; RNA_polII_repeat.
 DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
 DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
 DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
 DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
 DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
 DR Pfam; PF04992; RNA_pol_Rpb1_6; 1.
 DR Pfam; PF04990; RNA_pol_Rpb1_7; 1.
 DR Pfam; PF05001; RNA_pol_Rpb1_R; 29.
 DR SMART; SM00663; RPB1_N; 1.
 DR PROSITE; PS00115; RNA_POL_II_REPEAT; 43.
 KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
 KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
 FT ZN FING 71 87 C2H2-TYPE (POTENTIAL).
 FT DOMAIN 1590 1958 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
 FT CONFLICT 1067 1067 W -> L (IN REF. 2).
 FT CONFLICT 1449 1449 D -> Y (IN REF. 2).
 FT SEQUENCE 1970 AA; 217205 MW; 6876FC25692A657E CRC64;

Query Match 7.4%; Score 131; DB 1; Length 1970;
 Best Local Similarity 29.7%; Pred. No. 0.55;
 Matches 65; Conservative 23; Mismatches 91; Indels 40; Gaps 12;

QY 93 PLPGELAGTSRRVPSYSSONQSPISAFQSPISYQVSPSSSSP-PSPSRGEPPNNMSST 151
 DB 1764 PTSPSYSPSPNTP-TPSPNTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 1817
 QY 152 FPFPLRNGCIPSSLSRLISNSCPTPVSSPTSKNKP-----LPNMESIA-KQSM 202
 DB 1818 YTP-----SSPSYSPS---SPSYSPSPSKYPTSPSPSPSPSPSPSPSPSPSPSPSPSP 1869
 QY 203 AIAKQSMASFN-----PPYAVASAP-SPTHHQHTLTATPECCSDSSSTYDS 250
 DB 1870 TSPKYSPTSP 1929

```

0y      251 GHWISFOKFAQQCFASAMPJTSTFNUVXKPAQOMSPN 289
Db      1930 -VSPTSPKSGTSTGSPSTGVSPTSTFVLTSTPA---ISP 1964

RESULT 6
RPB1_MOUSE
ID      RPB1_MOUSE      STANDARD;      PRT; 1970 AA.
AC      P08775;
DT      01-NOV-1988 (Rel. 09, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      de-directed RNA polymerase II largest subunit (EC 2.7.7.6) (RPB1).
OS      Mus musculus (mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RN      SEQUENCE FROM N.A.
RA      MEDLINE=87280135; PubMed=3038894;
RX      Ahearn J.M., Jr., Bartolomei M.S., West M.L., Cieck L.J., Corden J.L.,
RT      "Cloning and sequence analysis of the mouse genomic locus encoding
RL      the largest subunit of RNA polymerase II.";
RL      J. Biol. Chem. 262:10695-10705 (1987).
RN      [2]
RN      SEQUENCE OF 1587-1970 FROM N.A.
RA      MEDLINE=86068017; PubMed=2999785;
RX      Corden J.L., Cadena D.L., Ahearn J.M., Jr., Dahmus M.E.;
RT      "A unique structure at the carboxyl terminus of the largest subunit
RL      of eukaryotic RNA polymerase II.";
RL      Proc. Natl. Acad. Sci. U.S.A. 82:7934-7938 (1985).
RN      [3]
RN      REVISIONS, AND PRESENCE OF AN ADDITIONAL EXON.
RX      MEDLINE=92178992; PubMed=1542581;
RA      Winzler M., Ackert J., Vicaite S., Vigneron M., Kedinger C.;
RT      "Complete sequence of the human RNA polymerase II largest subunit.";
RL      Nucleic Acids Res. 20:910-910 (1992).
CC      -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC      OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC      SUBSTRATES.
CC      -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC      {RNA} (N).
CC      -1- SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.
CC      -1- SUBCELLULAR LOCATION: Nuclear.
CC      -1- PTM: THE TANDDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC      -1- THE PHOSPHORYLATION ACTIVATES POL2.
CC      MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
CC      FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC      PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC      III FOR 5S AND TRNA GENES.
CC      -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC      at the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M12130; AAA40071.1;
DR      EMBL; M14190; AAA40071.1; JOINED.
DR      PIR; A28490; A28490.
DR      MCD; MGI:98086; Polr2a.
DR      GO; GO:0005634; C:nucleus;
DR      GO; GO:0003677; F:DNA binding activity; ISS.
DR      GO; GO:0003901; F:DNA-directed RNA polymerase II activity; ISS.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
DR      GO; GO:0006366; P:transcription from Pol II promoter; ISS.
DR      InterPro; IPR007022; RNA_pol_A.
DR      InterPro; IPR007080; RNA_pol_Rpb1_1.

```

DR InterPro: IPR007066; RNA_pol Rpb1_3.
 DR InterPro: IPR007083; RNA_pol Rpb1_4.
 DR InterPro: IPR007081; RNA_pol Rpb1_5.
 DR InterPro: IPR007075; RNA_pol Rpb1_6.
 DR InterPro: IPR007073; RNA_pol Rpb1_7.
 DR InterPro: IPR006592; RNA_pol N.
 DR InterPro: IPR006684; RNA_pol1_repeat.
 DR Pfam: PF04997; RNA_pol Rpb1_1.
 DR Pfam: PF00623; RNA_pol Rpb1_2.
 DR Pfam: PF04983; RNA_pol Rpb1_3.
 DR Pfam: PF05000; RNA_pol Rpb1_4.
 DR Pfam: PF04998; RNA_pol Rpb1_5.
 DR Pfam: PF04992; RNA_pol Rpb1_6.
 DR Pfam: PF04990; RNA_pol Rpb1_7.
 DR Pfam: PF05001; RNA_pol Rpb1_R.
 DR SMART; SM00663; RPOLA_N; 1.
 DR PROSITE; PS00115; RNA_POL_II_REPEAT; 42.
 KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
 DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
 FT FT 71 87
 FT DOKAIN 1590 1958 C2H-TYPE (POTENTIAL).
 FT CONFLICT 1498 1498 P -> R (IN REF. 1 AND 2).
 FT CONFLICT 1499 1536 MISSING (IN REF. 1 AND 2).
 FT SEQUENCE 1970 AA; 217175 MW; 70D6F38FD92A657E CRC64;
 SQ
 -Query Match 7.4%; Score 131; DB 1; Length 1970;
 Beac Local Similarity 29.7%; Pred. No. 0.55;
 Matches 65; Conservative % 23; Mismatches 91; Indels 40; Gaps 12
 QY 93 PLPGEIGTSSRLTPVSSQNSQPLISAFQSPISPYGVYSPSSSF-PSFSRREPNNMST 151
 Db 1764 PTSGSYSPSPNTPT-TSPNVSPSPSYSPSPSY--SPSPSYSPSSPRPTP--QSP 181
 QY 152 FFPPLRNGGIPSSLPILKINSCEVTPPVSSPTSKMPK-----LPNMESIA-KQSM 202
 Db 1818 YTP-----SSPSYSPS---SPSYSPSPKKTPTSPSPSPSPSPETTPASPKYSPSPKYS 186
 QY 203 AIAKQSMASNNY-----PFVAVASPA-SPTRHOFHTLATTPCEDSDSTVDS 250
 Db 1870 TSPKYSTPTPTPTPTPKYSPPTSPSPSPSPVPTPTSPKYSPTSPSPSPSPSPSPSPSP 1922
 QY 251 GHWISFQKFAQQQPFASNMVPTSPFTLVKPAQQNSPN 289
 Db 1930 -YSPTSPKGYSPPTSPGYSPPTSPYSLTSPA---ISPD 1964
 RESULT 7
 PIX2_XENLA STANDARD; PRT; 326 AA.
 ID PIX2_XENLA
 AC Q9PMR3; O93196; Q9YHA; Created
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pituitary homeobox 2 (Xp1x2).
 GN PITX2 OR PITX2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP MEDLINE=98370612; PUBMED=9707115;
 RA Ryan A.K., Blumberg B., Rodriguez-Berehan C., Yonet Tanura S.,
 RA Tamura K., Tsuchi T., de la Pena J., Sabbagh W., Greenwald J.,
 RA Choe S., Norris D.P., Robertson E.J., Evans R.M., Rosenfeld M.G.,
 RA Izpisua-Belmonte J.-C.;
 RT "Pitx2 determines left-right asymmetry of internal organs in
 vertebrates".
 RT Nature 394:545-551 (1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM PITX2A).
 RX MEDLINE=99146686; PUBMED=10021341;
 RX MEDLINE=99146686; PUBMED=10021341;


```

DR PROSITE / PS00771; HOMEOBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KM Alternative splicing;
FT DNA_BIND      85       144
FT DOMAIN        279     292
FT DOMAIN        285     289
FT FT            61
FT VASAPLIC      1
SO SEQUENCE      317 AA; 35320 MW; 188315708EBD95D CRC64;

Query Match          7.2% Score 126.5; DB 1; Length 317;
Best Local Similarity 25.9%; Pred. No. 0.14; Indels   59; Gaps   11
Matches    60; Conservative   25; Mismatches           88;

QY 6 ATSTSAATAAAAAAARAKRKPSPRENNRRERRARRAAKIYGLRAQGYNLPKHC DN 65
Db 119 AWTNLTLEAVRWVWFKRKARKRRERENQAELCKNGRGPG-FGLMGPIYDMIPGIS YN 177
QY 66 NEVLKALCVENGAVEEDGTYTRKGCKPLPGEIAGTSSRYTP-VSQNQSLPS--SAPGS 122
Db 178 NMAAKGL-----TSASTSKSPFPFNSNMVNPLLSSQSMTSP 213
QY 123 PISIVGVSPSSSFPPSRSEPN-----NMMSITFFPLRNCIGIPSLSPELRINSCEPV 176
Db 214 PNSISSMSSSWSPVASVTGTVPSSLNLNLNLNL6SPSL-NSAVPTP-----ACPY 264
QY 177 TPVSAPTSGKNPKPLPWESIAQQSMAT---AKOSMASFNYPYAAPSAPS 224
Db 265 APET-----PPYYRTCTNSSLASLRLKAQH-SFGVA--SVQNPA S 304

RESULT 11
ID PIX2 RAT STANDARD; PART; 324 AA.
AC OGRONH;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pituitary homeobox 2 {rpxx2}.
DS PITX2 OR PTX2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mamalia; Euarchonta; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_Taxid=10116;
(1) _SEQUENCE FROM N.A. (ISOFORM PTX2B).
RP STRAIN=Sprague-Dawley;
RC MEDLINE=98422376; PubMed=9748586;
RA Lindberg C., Wunderlich M., Ratliff J., Dinmore J.J., Jacoby D.B.;
RT "Regulated expression of the homeobox gene, rpx2, in the developing rat.";
RL Brain Res. Dev. Brain Res. 110:215-226(1998).
RN [2]
SB SEQUENCE FROM N.A. (ISOFORMS PTX2A AND PTX2B).
RX MEDLINE=20490920; PubMed=11032870;
RA Smidt M.P., Cox J.J., Van Schachck H.S.A., Coolen M., Scheepers J.,
RV van der Kleij A.M., Burbach J.P.H.;
RT "Analysis of three ptx2 splice variants on transcriptional activity and differential expression pattern in the brain.";
J Neurochem. 75:1818-1825(2000).
CC - FUNCTION: May play an important role in development and maintenance of anterior structures. May play a role in the proper localization of asymmetric organs such as the heart and stomach.
CC - SUBCELLULAR LOCATION: Nuclear.
CC - ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;

```


DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
 DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
 DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
 DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
 DR Pfam; PF04992; RNA_pol_Rpb1_6; 1.
 DR Pfam; PF04990; RNA_pol_Rpb1_7; 1.
 DR Pfam; PF05001; RNA_pol_Rpb1_R; 14.
 DR SMART; SMO0663; RPOA_N; 1.
 DR PROSITE; PS00115; RNA_POL_II_REPEAT; 22.
 DR Transferrase; DNA-directed RNA polymerase; Transcription; zinc; Repeat;
 KM DNA-binding; Nuclear protein; Phosphorylation; zinc-finger;
 KM 3D-structure.
 FT FT FING 67 83 C2H2-TYPE (POTENTIAL).
 FT DOMAIN 1544 1719 CARBOXYL-TERMINAL 7-RESDUE REPEATS.
 FT VARIANT 1553 1659 MISSING (IN STRAIN A364A).
 FT CONFLICT 1514 1514 A -> V (IN REF. 1).
 FT CONFLICT 1524 1524 G -> A (IN REF. 1).
 FT CONFLICT 1601 1601 T -> M (IN REF. 1).
 SQ SEQUENCE 1733 AA; 191610 MW; A45C1360FP99P68 CRC64;
 Query Match 7.2%; Score 126.5; DB 1; Length 1733;
 Best Local Similarity 28.1%; Pred. No. 0.94;
 Matches 64; Conservative 22; Mismatches 91; Indels 51; Gaps 10;
 QY 76 AGWVEEDGTYRKCKPL-----PGEIAGTSRTVPYSGNQSPISAFQSPIS 126
 DB 1505 AGGTATGAGAYGERTSPGAYGAPTSPG--FGVSGFGFPTSTPTSPATSPIS 1562
 QY 127 YQVSPSSSSF-----PSPRGEPNNMSTFPPEFLNNGIPSLPLRLSNSCPTVPVSS 182
 DB 1563 Y--SPTSPSYSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSP 1606
 QY 183 PTKPKPKLPWESIAKOSMAIAKOSMAFNVPFAVAPA-SPTRRHQFTLATIPCD 241
 DB 1607 PTS--PSYSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSP 1658
 QY 242 ESDSSTDSGHWISQKFAQQQPSASNVPTSPNTLVKPAQCMSPN 289
 DB 1659 PAYSPTSPS-----YSPTSPSYSPSTSPSTSPSTSPSTSPSTSPN 1695
 RESULT 13
 ID PIX2_CHICK STANDARD; PRT; 333 AA.
 AC 093385; 073854; 09PWE0;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pituitary homeobox 2 (PITX2).
 GN PITX2 OR PITX2.
 OS Gallus gallus (Chicken).
 OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCBI_Taxid=9031;
 RX MEDLINE=98300291; PubMed=9636662;
 RA St. Amant T.R., Ra J., Zhang Y., Hu Y., Baber S.I., Qiu M., Chen Y.;
 RT "Cloning and expression pattern of chicken Pitx2: a new component in
 the SHH signaling pathway controlling embryonic heart looping.";
 RL Biochem. Biophys. Res. Commun. 247:100-105(1998).
 RN (2)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Ryan A.K., Blumberg B., Rodriguez-Esteban C., Yonet-Tamura S.,
 RA Tamura K., Tsukitani T., de la Pena J., Sabbagh W., Greenwald J.,
 RA Choe S., Norris D.P., Robertson E.J., Evans R.M., Rosenfeld M.G.,
 RA Izpisua-Belmonte J.C.;
 RT "Pitx2 determines left-right asymmetries in vertebrates.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE OF 170-319 FROM N.A.

RA Albarjar M., Piedra M.E., Icardo J.M., Ros M.A., Rodriguez-Rey J.C.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: May play an important role in development and
 CC maintenance of anterior structures. May play a role in determining
 CC left-right asymmetry and in vasculogenesis during avian
 CC embryogenesis.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: Named isoforms=2;
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=O93385-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O93385-2; Sequence=VSP_002266; VSP_002267;
 CC -1- SIMILARITY: BELONGS TO THE PAIRED HOMEBOX FAMILY. BICOID
 CC SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 homeobox domain.
 CC -1- SIMILARITY: Contains 1 OAR domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC EMBL; AF076640; AAC27322.1; -
 CC EMBL; AF077092; AAD46097.1; -
 CC EMBL; AF063935; AAC18346.1; -
 CC PIR; J00097; J00097.
 CC HSP; P06601; IPI.L.
 CC InterPro; IPR003654; Homeo_OAR.
 CC InterPro; IPR003356; Homeobox.
 CC InterPro; IPR007104; Paired_homeo.
 CC Pfam; PF00046; homeobox; 1.
 CC Pfam; PF03826; OAR; 1.
 CC PRINTS; PR00024; HOMEBOX.
 CC ProDom; PD000010; Homeobox; 1.
 CC DR SMART; SMO0389; HOX; 1.
 CC DR PROSITE; PS00027; HOMEBOX_1; 1.
 CC DR PROSITE; PS00071; HOMEBOX_2; 1.
 CC DR PROSITE; PS0803; OAR; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 CC Alternative splicing.
 CC FT DNA_BIND 101 160 HOMEBOX.
 CC FT DOMAIN 295 308 OAR.
 CC FT DOMAIN 301 305 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DOMAIN 22 33 POLY-SER.
 CC FT VARSPLIC 1 62 Missing (in isoform 2).
 CC FT VARSPLIC 63 77 LEVHTISPTSPPEAA -> MESNCRKLVSACVQL (in
 CC isoform 2).
 CC FT VARSPLIC 63 77 /FtId=VSP_002267.
 CC FT CONFLICT 234 234 S -> T (IN REF. 3).
 CC FT CONFLICT 292 292 T -> M (IN REF. 3).
 CC FT CONFLICT 317 317 N -> K (IN REF. 3).
 CC FT CONFLICT 328 328 P -> A (IN REF. 2).
 SQ SEQUENCE 333 AA; 36557 MW; FE9C68E475F5C85 CRC64;
 Query Match 7.1%; Score 125.5; DB 1; Length 333;
 Best Local Similarity 25.9%; Pred. No. 0.17;
 Matches 60; Conservative 27; Mismatches 86; Indels 59; Gaps 11;
 QY 6 ATSTSAAAAAAAAAARRRKPSRRERRRRRAVAATYGLRAGQDYNLAKHCN 65
 DB 135 AVNTNLTEARVRVFKRRRAKRRERNOALCNGGPGP-FNGLMOPYDMDYGVSYN 193
 QY 66 NEVLKALCVKAGWVEEDGTYRKCKPLPEIAGTSRTVP-YSSQNSPIS--SAROS 122
 DB 194 NMAKGL-----TSASLSTKSPFPFNSMNVNPLSSQSMFSP 229
 QY 123 PIRYQVSPSSSSFPSPSRGEPN-----NMSTTFPEFLNNGIPSLPLRLSNSCPV 176

DB 230 PUSISMSMSMSVPAATVGVPCSGSLNSLNLNN-----LSNPSLNSAYP-----TPACEY 280

QY 177 TTPVSSPTSKNPKPLPNMESIAKQSMAI-----AKQSMASENFPYVAVSAPAS 224

DB 281 ABPT-----PPYVYDTCNSLSLALRLKAKQH-SSRQVYA--SVQNPAS 320

RESULT 14

REB1 DROME STANDARD; PRT; 1887 AA.

AC P04052; OSVYX5;

DT 01-NOV-1986 (Rel. 03, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE DNA-directed RNA polymerase II (last annotation update)

GN RPII215 OR CG1554.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OC NCBI_Taxid=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89218930; PubMed=2496296;

RT Jokeit R.S., Weeks J.R., Zehring W.A., Greenleaf A.L.;

RT "Analysis of the gene encoding the largest subunit of RNA polymerase II in Drosophila."

RT Mol. Gen. Genet. 215:266-275 (1989).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=Beirley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Ball R.J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Boltenkov S.,

RA Borokva D., Botchan M.R., Bouck J., Brodeur P., Brothier P.,

RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Casley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,

RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hestlin D., Houson K.A., Howland T.J., Wei M.-H., Ibegwan C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Matcel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Stizkeas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng H.O.,

RA Gbba R.A., Myers B.W., Rubin G.M., Venter J.C.;

RT "the genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195 (2000).

RN [3]

RP SEQUENCE OF 1-27 FROM N.A.

RX MEDLINE=87089662; PubMed=3025586;

RA Searles L.L., Greenleaf A.L., Kemp W.E., Voelker R.A.;

RT "Sites of P element insertion and structures of P element deletions in

RT the 5' region of Drosophila melanogaster RpiI215.";

RL Mol. Cell. Biol. 6:3312-3319 (1986).

RN [4]

RP SEQUENCE OF 1-472 FROM N.A.

RX MEDLINE=85282618; PubMed=2992806;

RA Biggs J., Searles L.L., Greenleaf A.L.;

RT "Structure of the eukaryotic transcription apparatus: features of the

RT gene for the largest subunit of Drosophila RNA polymerase II.";

RL Cell 42:611-621 (1985).

RN [5]

RP SEQUENCE OF 1441-1887 FROM N.A.

RX MEDLINE=88094402; PubMed=3122024;

RA Allison L.A., Wong J.K.-C., Fitzpatrick V.D., Moyle W., Ingles C.J.;

RT "The C-terminal domain of the largest subunit of RNA polymerase II of

RT Saccharomyces cerevisiae, Drosophila melanogaster, and mammals: a

RT conserved structure with an essential function.";

RL Mol. Cell. Biol. 8:321-329 (1988).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS

CC SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC (RNA) (N).

CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO

CC TWELVE DIFFERENT POLYPEPTIDES. THE 215 KDA POLYPEPTIDE IS THE

CC LARGEST COMPONENT OF RNA POLYMERASE II.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- PTM: THE TANDEN 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.

CC -1- THE PHOSPHORYLATION ACTIVATES POL2.

CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE

CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA

CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE

CC III FOR 5S AND TRNA GENES.

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; M27431; AAA28868.1; -

CC EMBL; AE003486; AAF8057.1; -

CC EMBL; M14203; AAA28864.1; -

CC EMBL; M1798; AAA28863.1; -

CC EMBL; M19537; AAA28827.1; -

CC PIR; S04457; RNPF2L.

CC FLYBase; FBgn0003277; RPII215.

CC InterPro: IPR000722; RNA_POL_A.

CC InterPro: IPR007080; RNA_POL_Rpbl_1.

CC InterPro: IPR007066; RNA_POL_Rpbl_3.

CC InterPro: IPR007083; RNA_POL_Rpbl_4.

CC InterPro: IPR007081; RNA_POL_Rpbl_5.

CC InterPro: IPR007075; RNA_POL_Rpbl_6.

CC InterPro: IPR007073; RNA_POL_Rpbl_7.

CC InterPro: IPR006592; RNA_POL_N.

CC InterPro: IPR000684; RNA_POLII_repeat.

CC Pfam; PF04997; RNA_pcl_Rpbl_1; 1.

CC Pfam; PF00623; RNA_pcl_Rpbl_2; 1.

CC Pfam; PF04983; RNA_pcl_Rpbl_3; 1.

CC Pfam; PF05000; RNA_pcl_Rpbl_4; 1.

CC Pfam; PF04998; RNA_pcl_Rpbl_5; 1.

CC Pfam; PF04992; RNA_pcl_Rpbl_6; 1.

CC Pfam; PF04990; RNA_pcl_Rpbl_7; 1.

CC Pfam; PF05001; RNA_pcl_Rpbl_R; 22.

CC SMART; SM00663; RPII215; 1.

CC PROSITE; PS00115; RNA_POL_II_REPEAT; 11.

CC Transferrase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;

CC DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.

CC ZN_FING 67 83 C2H2-TYPE (POTENTIAL).

FT DOMAIN 1579 1881 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
 FT CONFLICT 319 324 RAMOKS -> GYAKV (IN REF. 4).
 FT CONFLICT 450 450 F -> G (IN REF. 4).
 FT CONFLICT 455 458 TLHK -> RCTT (IN REF. 4).
 FT CONFLICT 463 472 GHRVYLPWS -> VTGESVAST (IN REF. 4).
 FT CONFLICT 741 741 R -> H (IN REF. 1).
 FT CONFLICT 1485 1524 SMIGAMFVGSGTSPMTPTMNCNTPRFSPGHV
 FT CONFLICT 1506 1508 MTP -> ELDSDA (IN REF. 1).
 FT CONFLICT 1887 1887 D -> DVKGRG (IN REF. 1).
 SQ SEQUENCE 1887 AA: 209167 MW: 4EC68C7708A167A3 CRC64;

Query Match 7.1%; Score 124.5; DB 1; Length 1887;
 Best Local Similarity 27.1%; Pred. No. 1.4;
 Matches 68; Conservative 16; Mismatches 78; Indels 89; Gaps 14;

QY 95 PGEIAGTSRVTPTYSQNGSPPLSARQSP-----IPSYQ-----VSPSSSSSF----- 136
 DB 1591 PG---GASPNVSP---SSPNVSPSPVLAADRYASTTNTFNPQSTGVSPPSSGVSPTSPVYS 1646
 QY 137 -----PSPSRGEPNNNMSTPTFPPLRNGGIFSSLPGLRI 170
 DB 1647 PTVOFQSSPPSPAGSSNIVSPGNAYSPSSSNVSPSPSPSPSPSPSPSPSPSPSPSPSPSP 1701
 QY 171 SNSC-PVTPPVSSPTSPKPKPL-----PNMESIAKQSMALAKQSMASFPYPPYAVSA 221
 DB 1702 TSPCTSP 1755
 QY 222 PASPTHRHQFTLATIPEDCESDSS---TVDSGHMISFQKFAQOQPFSAWMTPTFNL 278
 DB 1756 PTPSPYS-----PSPSPVDSFGSPQYTPGS-----FOYSPASXKYSPTSP--L 1797
 QY 279 VKPAPQMSPN 289
 DB 1798 YSP 1808

RESULT 15
 AMYH YEAST
 AC P08640; P08068; STANDARD; PRT; 1367 AA.
 DT 01-AUG-1988 (Rel. 08. Created)
 DT 01-FEB-1995 (Rel. 31. Last sequence update)
 DT 15-SEP-2003 (Rel. 42. Last annotation update)
 DE Glucosylase SI/S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucosylase).
 GN STAI OR STAZ OR MAL5 OR YIR019C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes,
 OC Saccharomycetales, Saccharomycetaceae, Saccharomycetes.
 NCBI TaxID=4932;
 RX STRAIN=5288C / AB972;
 RC PubMed=9169870;
 RA Churcher C.W., Bowman S., Badcock K., Bankier A., Brown D.,
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
 RA Harris D.E., Horenell T., Hunt S., Jagsels K., Jones M., Lye G.,
 RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.,
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.",
 Nature 387:84-87(1997).
 RL [2]
 RP SEQUENCE OF 1-242 AND 762-1331 FROM N.A.
 RX MEDLINE=87194600; PubMed=3106330;
 RA Yamashita I., Nakamura M., Fukui S.;
 RT "Gene fusion is a possible mechanism underlying the evolution of
 RT STR1.",
 RL J. Bacteriol. 169:2142-2149 (1987).
 RP SEQUENCE OF 1-31 FROM N.A.
 RC STRAIN=SPY101-1C;
 RX MEDLINE=89031230; PubMed=3141213;

RA Pardo J.M., Ianez E., Zalacain M., Claros M.S., Jimenez A.;
 RT "Similar short elements in the 5' regions of the STA2 and SGA genes
 RT from Saccharomyces cerevisiae.",
 RL FEBS Lett. 239:179-184 (1988).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SIMILARITY: TO S. POMBE SPC215.13.
 CC -1- SIMILARITY: SOME. TO S. POMBE SPC285.13C.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 DR EMBL: Z38061; CA86176.1; -
 DR EMBL: M16164; AA35014.1; -
 DR EMBL: M16165; AA35015.1; -
 DR EMBL: X13857; CA32069.1; -
 DR PIR: S48478; S48478.
 DR SGD: S0001458; MUC1.
 DR GO: GO:0005886; C:plasma membrane; IDA.
 DR GO: GO:0030447; P:filamentous growth; IDA.
 DR GO: GO:0007125; P:invasive growth; IMP.
 DR GO: GO:0007124; P:pseudohyphal growth; IMP.
 KW Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KW signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 1367
 FT DOMAIN 210 1367
 FT CARBOHYD 817 817
 FT CARBOHYD 874 874
 SQ SEQUENCE 1367 AA: 136110 MW: 91C0DE2DBD61AA9D CRC64;

Query Match 7.0%; Score 124; DB 1; Length 1367;
 Best Local Similarity 27.1%; Pred. No. 1;
 Matches 64; Conservative 35; Mismatches 105; Indels 32; Gaps 12;

QY 85 TTYRCKPLDGEIAGT---SSRVTPYSQNGSPPLSARQSPDIPSYQVSPSSSPSPSPSP 141
 DB 698 STTESSAPVPTPSSSTTESSAPVPTPSSSTTESSA---FVP---TPSSSTTESSSA 750
 QY 142 GEPNNMSTFPFLRNGGIFSSLPGLRIASNCPTTPPVSSPTSPKPKPLN-WEIAKO 200
 DB 751 PVTSTTESSAE-----VTPSSSTTESSAPVPTPSSSTTESSAPVPTPSSSTTES 804
 QY 201 SMAIAKQSMASFPYPPYAVSAPSPTRHQFTLATIPEDCESDSTVDSGHMISFQKFA 260
 DB 805 SVAPVPTPSSSSN---ITSSAPSTPSSSTTES-SSVVPPTPSSSTTESSAPVS---S 856
 QY 261 QQQPSPASWPT--SPTFNLVKPAPQO---SPNTAFOEIQGS---SEPKFENSQVK 310
 DB 857 STTESSVAPVPTPSSSSNITSSAPSPDIPSPSTTEF-STGTVTPTPSSSKYSPESQTE 911

Search completed: December 19, 2003, 14:27:05
 Job time : 19 secs

Db 132 NMAAKGL-----TSASLSTKSPFPNSMNVNPLSSQSMSP 167
Qy 123 PIPSVQVSPSSSPSPSRGEPN-----NNMSSTFFPLRNGGIPSLPLRLISNCPV 176
Db 168 PNISSMSMSSMWPASATGVPGSSLNINLNINLNLSSPSL-NSAVPTP-----ACPY 218
Qy 177 TPPVSSPTSKPKPLPNMESIAKQSMAL-----AKQSMASFNYPFYAVSAPAS 224
Db 219 APPT-----PPYVYRDTCSNLSLRLAKQH-SSFGYA--SVQKPAS 258

RESULT 2

US-08-754-477A-121
Sequence 121, Application US/08754477A
Patent No. 6518411
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477A
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-477A-121

Query Match 7.2%, Score 127.5; DB 4; Length 271;
Best Local Similarity 25.9%, Pred. No. 0.00067;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;
Qy 6 ATSTSAIAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGIRAGQDYNLRKHCN 65
Db 73 AVMTLLEARYVWFKNRAKRRKERNQOAEICNGGPGQ-FNGLMGPYDMDYGYSYN 131
Qy 66 NEVLKALCEAGWVEEDGTTYRKCKPLPEIATGSSRVT-YSQNSQPLS--SAFOS 122
Db 132 NMAAKGL-----TSASLSTKSPFPNSMNVNPLSSQSMSP 167
Qy 123 PIPSVQVSPSSSPSPSRGEPN-----NNMSSTFFPLRNGGIPSLPLRLISNCPV 176
Db 168 PNISSMSMSSMWPASATGVPGSSLNINLNINLNLSSPSL-NSAVPTP-----ACPY 218
Qy 177 TPPVSSPTSKPKPLPNMESIAKQSMAL-----AKQSMASFNYPFYAVSAPAS 224
Db 219 APPT-----PPYVYRDTCSNLSLRLAKQH-SSFGYA--SVQKPAS 258

RESULT 3

US-08-754-477A-125
Sequence 125, Application US/08754477A
Patent No. 6518411
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477A
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-477A-125

Query Match 7.2%, Score 127.5; DB 4; Length 271;
Best Local Similarity 25.9%, Pred. No. 0.00067;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

Qy 6 ATSTSAIAAAAAAAAAARRKPSWRENNRRRRRRRAVAAKIYTGIRAGQDYNLRKHCN 65
Db 73 AVMTLLEARYVWFKNRAKRRKERNQOAEICNGGPGQ-FNGLMGPYDMDYGYSYN 131
Qy 66 NEVLKALCEAGWVEEDGTTYRKCKPLPEIATGSSRVT-YSQNSQPLS--SAFOS 122
Db 132 NMAAKGL-----TSASLSTKSPFPNSMNVNPLSSQSMSP 167
Qy 123 PIPSVQVSPSSSPSPSRGEPN-----NNMSSTFFPLRNGGIPSLPLRLISNCPV 176
Db 168 PNISSMSMSSMWPASATGVPGSSLNINLNINLNLSSPSL-NSAVPTP-----ACPY 218
Qy 177 TPPVSSPTSKPKPLPNMESIAKQSMAL-----AKQSMASFNYPFYAVSAPAS 224
Db 219 APPT-----PPYVYRDTCSNLSLRLAKQH-SSFGYA--SVQKPAS 258

RESULT 4
US-08-957-351-26
Sequence 26, Application US/08957351
Patent No. 6306586
GENERAL INFORMATION:
APPLICANT: Semina, Elena
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF CATARACTS
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,351
FILING DATE: 24-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-024.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-957-351-26

Query Match 7.2%; Score 126.5; DB 4; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.00083;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTSAIAAAARRRKPSWRENNRRRRRAVAKITYTGLRAQGDYVLPKHCND 65
DB 73 AWMITLTVARVYWFKNRRAKWKRRNQAEICKXGFGPQ-FNGLMQPYDDMYPGYSYN 131
QY 66 NEVLKALCVKAGVVEEDGTYRKCKPLPGEIAGTSSRVTP-YSSONOSPLS-SAFQS 122
DB 132 NMAAKGL-----TSASLSTKSPFPFNSMNVNPLSSQSMFSP 167
QY 123 PIPSYOVSPSSSPSPSRGEPN-----NNMSSTFPPLRNGIPLSLPSLRISNSCPV 176
DB 168 PMSISMSSMSMVSATVGVGSSLSLNNLNNLSSPFL-NSAVPTP-----ACPY 218
QY 177 TPVSSPTSKNPKPLPNMESIAKQSMAT---AKQSMASFNPFYVASAPAS 224
DB 219 APPT-----PPYVYRDTGNCSSLSLAKAKQH-SSFGYA--SVQNPAS 258

RESULT 5
US-08-754-477A-5
Sequence 5, Application US/08754477A
Patent No. 6518411
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
APPLICANT: Semina, Elena
TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
NUMBER OF SEQUENCES: AND DIAGNOSTIC USES THEREFOR
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/754,477A
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-477A-5

Query Match 7.2%; Score 126.5; DB 4; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.00083;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTSAIAAAARRRKPSWRENNRRRRRAVAKITYTGLRAQGDYVLPKHCND 65
DB 73 AWMITLTVARVYWFKNRRAKWKRRNQAEICKXGFGPQ-FNGLMQPYDDMYPGYSYN 131
QY 66 NEVLKALCVKAGVVEEDGTYRKCKPLPGEIAGTSSRVTP-YSSONOSPLS-SAFQS 122
DB 132 NMAAKGL-----TSASLSTKSPFPFNSMNVNPLSSQSMFSP 167
QY 123 PIPSYOVSPSSSPSPSRGEPN-----NNMSSTFPPLRNGIPLSLPSLRISNSCPV 176
DB 168 PMSISMSSMSMVSATVGVGSSLSLNNLNNLSSPFL-NSAVPTP-----ACPY 218
QY 177 TPVSSPTSKNPKPLPNMESIAKQSMAT---AKQSMASFNPFYVASAPAS 224
DB 219 APPT-----PPYVYRDTGNCSSLSLAKAKQH-SSFGYA--SVQNPAS 258

RESULT 6
US-08-754-477A-131

Sequence 131, Application US/08754477A
Patent No. 6518411

GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
APPLICANT: Semina, Elena
TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
NUMBER OF SEQUENCES: AND DIAGNOSTIC USES THEREFOR
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477A
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 131:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-477A-131

Query Match 7.1% Score 124.5; DB 4; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.0013;
Matches 60; Conservative 24; Mismatches 89; Indels 59; Gaps 11;

QY 6 ATSTSAAMAAAAAARRKPSWRENNRRRRRAVAAKIYTLGRAQGDYNLPHKCDN 65
DB 73 AVMTLTBARVWPKRPAKWKRRKRNQOALCKNGGCPQ-FNGLMQPYDDMYGYGSYN 131
QY 66 NEVLKALCVEAGWVVEEDGTTTKGCKPLPGEIAGTSRATP-YSSONQSPIS--SARQS 122
DB 132 NMAAKGL-----TSASLSTKSPFPFNSMNVNPLSSQSMFSP 167
QY 123 PIPSVQVSSSPSSPSPSRGEPN-----NNMSSTFPPLRNGGIPSLPLRISNSCPV 176
DB 168 PMSISMSMSMWSVPSAVTGVPGSSLNLSLNLNLNLSBSL-NSAVPTP-----ACPY 218
QY 177 TPVSSPTSKPKPLPNMESIAKQSMAL---AKQSMASFNPYAVSNAPAS 224
DB 219 APPT-----PPYVYRDTGNSSLASLRKAKQH-SSEFGYA--SVQKRPAS 258

RESULT 7
US-08-728-323A-2
Sequence 2, Application US/08728323A
Patent No. 5948676

GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Immediate Early Protein From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
TITLE OF INVENTION: Encoding Same And Uses Thereof
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSER: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-728-323A-2

Query Match 7.0% Score 124; DB 2; Length 1162;
Best Local Similarity 22.5%; Pred. No. 0.012;

Matches 79; Conservative 49; Mismatches 129; Indels 94; Gaps 17;

QY 8 STSAAAAAARRKPSWRENNRRRRRAVAAKIYTLGRAQGDYNLPHKCDNNE 67
DB 13 STGAPLTRGSCRKRRNSPERCDLGLHLPKRRKVAOSI--DGRCCGPHLTP----- 63
QY 68 VLKALCVEAGWVVEEDGTTTKGCKPLPGEIAGTSRATP-YSSONQSPIS--PLSS 118
DB 64 -----LPGSFTVFTSG---LPAFVSSPLPAPLPSPAPATLPPLPALLPPT 108
QY 119 AFOSPI-PSYQVSSSPSSPSPSRG-BENNMSSTFPPLRNGGIPSLPLRISNSCPV 176
DB 109 TSSSPIPSHVSQGTDTHTSPSPALPPTQSPESQRP-----PLSSPTGRPDSTPM 161
QY 177 TPVVS-----SPTSKPKP--LPNMESIAKQSM-ALAKQSMASFNPYF-----YAVSNP 222
DB 162 RPPSQQTTPHSPSTTPPEPPSKSPDGLPSTLRSLRKRLSSPQGSTLNPICQSP 221
QY 223 ASPTRHQHTLATIPEDCESDSTVDGHWISFQKPAQOQPF---SAS-----MVPT 272
DB 222 VSP-----PRCDPANRSV-----YPMATESPIYVGSSSDGDTPPROPPT 261
QY 273 SP-TFNLYKPAPOQSPNTAFAF-----OEIGQSEKFEFNSQVQKPEGE 315
DB 262 SPIGSSSPSEGSWGDDTAMLVLAETAEASXKXKCEENQAGEDWDG 312

RESULT 8
US-09-298-568-2
Sequence 2, Application US/09298568
Patent No. 6322792

GENERAL INFORMATION:
APPLICANT: Kieft, Elliott D.
APPLICANT: Ballester, Mary E.
APPLICANT: Kaye, Kenneth M.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
TITLE OF INVENTION: RHADINO VIRUS DNA TO MEDIATE EFFICIENT EBIOSOME PERSISTENCE
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/298,568
CURRENT FILING DATE: 1999-04-21
EARLIER APPLICATION NUMBER: US 60/109,422
EARLIER FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patent In Ver: 2.0
SEQ ID NO 2
LENGTH: 1162
TYPE: PRT
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-2

Query Match 7.0% Score 124; DB 4; Length 1162;
Best Local Similarity 22.5%; Pred. No. 0.012;
Matches 79; Conservative 49; Mismatches 129; Indels 94; Gaps 17;

QY 8 STSAAAAAARRKPSWRENNRRRRRAVAAKIYTLGRAQGDYNLPHKCDNNE 67
DB 13 STGAPLTRGSCRKRRNSPERCDLGLHLPKRRKVAOSI--DGRCCGPHLTP----- 63
QY 68 VLKALCVEAGWVVEEDGTTTKGCKPLPGEIAGTSRATP-YSSONQSPIS--PLSS 118
DB 64 -----IPGSPYFTSG---LPAFVSSPLPAPLPSPAPATLPPLPALLPPT 108
QY 119 AFOSPI-PSYQVSSSPSSPSPSRG-EPNNMSSTFPPLRNGGIPSLPLRISNSCPV 176
DB 109 TSSSPIPSHVSQGTDTHTSPSPALPPTQSPESQRP-----PLSSPTGRPDSTPM 161
QY 177 TPVVS-----SPTSKPKP--LPNMESIAKQSM-ALAKQSMASFNPYF-----YAVSNP 222
DB 162 RPPSQQTTPHSPSTTPPEPPSKSPDGLPSTLRSLRKRLSSPQGSTLNPICQSP 221
QY 223 ASPTRHQHTLATIPEDCESDSTVDGHWISFQKPAQOQPF---SAS-----MVPT 272
DB 222 VSP-----PRCDPANRSV-----YPMATESPIYVGSSSDGDTPPROPPT 261

Qy 273 SP-TFNLVKPAPQOMSPTAF-----OEIGSSSEFKENSQVPMGE 315
 Db 262 SPISISSSSSESGWGDPTMLVLAEIAEASKEKESENNQAGEDND 312

RESULT 9
 US-09-410-399-2
 Sequence 2, Application US/09410399
 Patent No. 6482587
 GENERAL INFORMATION:
 APPLICANT: Robertson, Erle S.
 APPLICANT: Colter, Murray A.
 TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
 FILE REFERENCE: UM-03778
 CURRENT APPLICATION NUMBER: US/09/410,399
 CURRENT FILING DATE: 1999-10-01
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 2
 LENGTH: 1162
 TYPE: PRT
 ORGANISM: Kaposi's sarcoma-associated herpesvirus
 US-09-410-399-2

Query Match 7.0%; Score 124; DB 4; Length 1162;
 Best Local Similarity 22.5%; Pred. No. 0.012;
 Matches 79; Conservative 49; Mismatches 129; Indels 94; Gaps 17;

Qy 8 STSAAAAAARRRRRRRRRRRRRRAVAAKIYTLBAQGDYVLPKHDNNE 67
 Db 13 STGAPLTRGSCRRKNSPERCDLGDHLQPRKHVADSL-DGRGCPHTLP----- 63
 Qy 68 VLKALCVAAGVVEEDGTYRKCKPLPGIAGTSRYTPYSSQNS-----PLSS 118
 Db 64 -----IPGSPVTFSG--LPAFVSSPTLPVAPIPBPATPLPPALLPPVT 108
 Qy 119 AFOSPT-PSYQVSPSSSPSPSPSG-EPNNNSSTFPPLRNGIISLSLRISNCPV 176
 Db 109 TSSSPISPEHPVSPGTTDHSFSPALPPTQSPSSQRP-----PLSPTGRPSSTM 161
 Qy 177 TRPV-----SPTSKNKP--LPMESIAQSM-AIAKQSMASNTYF---YAVSAP 222
 Db 162 RPPPOQTTPHSPPTPPEPPSKSPDLAPETLSLKRKRSPPQGSTLNPICOSP 221
 Qy 223 ASPTRHQFTLATIPECDESSTVDSGHWISFOKFAOQOPF---SAS-----MPT 272
 Db 222 VSP-----PRCDPANRSV-----YPMATESPIYVGSSSDGDTPPROPPT 261
 Qy 273 SP-TFNLVKPAPQOMSPTAF-----OEIGSSSEFKENSQVPMGE 315
 Db 262 SPISISSSSSESGWGDPTMLVLAEIAEASKEKESENNQAGEDND 312

RESULT 10
 US-08-227-536-2
 Sequence 2, Application US/08227536
 Patent No. 5658784
 GENERAL INFORMATION:
 APPLICANT: Eckner, Richard
 APPLICANT: Ewen, Mark
 APPLICANT: Livingston, David
 TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESS: Weingarten, Schurgin, Gagnebin & Hayes
 STREET: Ten Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/227,536
 FILING DATE: 14-APR-1994
 CLASSIFICATION: 436
 ATTORNEY/AGENT INFORMATION:
 NAME: Williams Ph.D., Kathleen A.
 REGISTRATION NUMBER: 34,380
 REFERENCE/DOCKET NUMBER: DPCI-308XX
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 451-0313
 TELEFAX: (617) 542-2290
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2414 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-227-536-2

Query Match 6.9%; Score 122.5; DB 1; Length 2414;
 Best Local Similarity 21.0%; Pred. No. 0.047;
 Matches 94; Conservative 35; Mismatches 166; Indels 153; Gaps 15;

Qy 2 TSDGARSTAAABAAAAAARRRRRRRRRRRRRRAVAAKIYTLBAQGDYVLPK 61
 Db 594 TDPALAKDRMENVLAAYARKVEGDYVESANNRAE-----YHLLAKIKYIKQ 642
 Qy 62 HCDNN-----EVKALCVAAGV-----VEEDGTYRKCKPLPGIAGT----- 101
 Db 643 ELEBKRTQLQONMLPNAAGVPMVPMGPMGQPGHT-SNGPLPDSMTIRGSVPNQ 701
 Qy 102 -SSRYTPSSQNSQPLSSAFQSPISYQV-----SPSSS 134
 Db 702 KMPRIITPQGLNQFGQMSMAQPIVPRQRPPIQHQQALQPGALNPMGCGPMQPSNQ 761
 Qy 135 SFPSSRGEPNNMSTFPPLRNGIISLSLRISNCPVTPVPSPTSK----- 186
 Db 763 GQFLPQTOPFSGQMVNIPILAPSSGQAFVSOAQWSSSCPVNPIIMPSSQSHIHCQ 821
 Qy 187 -----NPKPLNM-----ESIAKQ-----SMIAKSMASFN 214
 Db 822 LPQALHONSPEPVSRITPHTTPPSIGAQQPATTIPAVPTPPAMPGPQSOALHP 881
 Qy 215 PFYAVSAPASPTERHQFTLATIPECDESSTVDSGHWISFOKFAOQOPF---ASW 270
 Db 882 PQGTPTPTTQLPQOVQSLAPASADQ-----FOQPRSQOSTAASV 924
 Qy 271 PFSPTFNLVKPAP-----QQNSP-----NTAFOEIGSSSEFKFE--- 305
 Db 925 PT-PNAPLPPQAPATPLSQPAVSIQGVSNPSTSTEVNSQALAEKQSOEYKQAKME 983
 Qy 306 -----NSQVPMGERIHVGMEDLE 326
 Db 984 VDQPEADTQPEDISEKVEDCKMESTE 1011

RESULT 11
 PCT-US95-04682-2
 Sequence 2, Application PCT/US9504682
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESS: Weingarten, Schurgin, Gagnebin & Hayes
 STREET: Ten Post Office Square
 CITY: Boston

STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-April-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-308Xq999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 451-0313
FAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04682-2

Query Match: 6.9%; Score 122.5; DB 5; Length 2414;
Best Local Similarity 21.0%; Pred. No. 0.047;
Matches 94; Conservative 35; Mismatches 166; Indels 153; Gaps 15;

QY 2 TSDGATSTSAIAAAAAAAAAARKEPWRERENRRRRRAVAAKITGTPRAQGDYNLPK 61
DB 594 TPDPAALXDRMENLVAAARKEGDMYBSANNRA-----YHLLAEKIKYK 642
QY 62 HCDNN-----EVLKALCVEAGWV-----VEDGTYRKCKPLPGIAGT----- 101
DB 643 ELEBGRTRRLQKONLPRNAAGVVPVSMNPGMGPQGPQMT-SNGPLPSPMIRSVPMQ 701
QY 102 -SSRVTPYSSONQSPSSAFQSPISYQV-----SPSSS 134
DB 702 MMRPRTPGSLNQFGQMSWAPPIVRQTPRLQHHGQLAQLALNPMGCGPRMQGPSNQ 761
QY 135 SPSSSRGEPNNMSTFPFLRNGISPSLSLRISNCPYTPPVSSFTSK----- 186
DB 762 GQFLPOTQFPQSGMVTNPLAPSSQAFVSOQMSSSSCPVNSPIMPGSQGSHHCPO 821
QY 187 -----NPKPLPNW-----ESIAKO-----SMALAQSMASFMV 214
DB 822 LPQALHONSPSPVSRTPPHHTPPSICAQOPRATTIPAVPTPPAMPFGQSOALHP 881
QY 215 PFYVASAPSPTRHIOFTLATIPEDSDSSTVDSGHWISQKFAQCPFS-----ASNV 270
DB 882 PROGTPTPTTQLPQOVQPSLPAFASDQ-----PQQQPRSGQSTRAASV 924
QY 271 PTPSPFNLVKAP-----QQNSP-----NTAAFOIGSSSEKFE----- 305
DB 925 PTPAPPLPPOAPATPLSOPAVSIEQVSNPSTSTEVNSQALIEKPSOEKQKAKME 983
QY 306 -----NSQVAPWESGERHIDVGMEDLE 326
DB 984 VDOPEPADTQPEDISKVEDCKMESTE 1011

RESULT 12
US-09-423-890-8
Sequence 8, Application US/09423890
Patent No. 6312934
GENERAL INFORMATION:

APPLICANT: CADUS PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: HUMAN MEKX PROTEIN AND NUCLEIC ACID MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: CPT-085CPC
CURRENT APPLICATION NUMBER: US/09/423,890
CURRENT FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: USSN 60/076,153
PRIOR FILING DATE: 1998-03-16
PRIOR APPLICATION NUMBER: USSN 60/099,165
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 8
LENGTH: 1493
TYPE: PRT
ORGANISM: Mus musculus
US-09-423-890-8

Query Match: 6.8%; Score 120.5; DB 4; Length 1493;
Best Local Similarity 22.2%; Pred. No. 0.035;
Matches 76; Conservative 36; Mismatches 131; Indels 99; Gaps 12;

QY 5 GATSTSAIAAAAAAAAAARKEPWRERENRRRRRAVAAKITGTPRAQGDYNLPK 64
DB 139 GATSPAGAEPPSA-----PSGREMEK-----ETLGLHMKED--REERN 179
QY 65 NNEVLKALCVEAGWVEEDGTYRK-----CKPLPEIAGTSRVTPYSSONQSPSSAFQ 121
DB 180 IREKLKATCMPA-WHEWLERRRNRGRPVVVPDIP--IKGDGSEVNNLAEPQEGAGSA 236
QY 122 SPIPSYQVSPSSSPSSRGEPPNNMSTFPFLRNGISPSLSLRISNCPV----- 176
DB 237 AAPAGRSPSPGSSPSGRSVKPS-----DVRKRKRVSPVFPQSG 277
QY 177 -TPVSSPTSKNPKPL-PNWEISIAQMAIAKQSMASFNYPYVASAPSPTR----- 228
DB 278 RITPRRAPSPDGFSPYSPETSRRVKNYMRARLLQOIGNSFLIGSDSDNKYRVFI 337
QY 229 -----HQFTLATIPEDSD-----STVDS 250
DB 338 GPONSCGGAFCIHLFVMLRVFQLEPSDPMILMKTLKNFEVESLFOKHSRRSSRIKA 397
QY 251 GHWISFOKE-----AQQPFASMTPTSPFNLVKAPQOMSP 288
DB 398 PSRNTIQKVRMSNHTLSSSTSTSSSENSIKOEBEOWCP 439

RESULT 13
US-08-628-829-4
Sequence 4, Application US/08628829A
Patent No. 6333170
GENERAL INFORMATION:
APPLICANT: Johnson, Gary L.
TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Extern
FILE REFERENCE: CPT-004DVC3
CURRENT APPLICATION NUMBER: US/08/628,829A
CURRENT FILING DATE: 1996-04-05
EARLIER APPLICATION NUMBER: 08/440,421
EARLIER FILING DATE: 1995-05-15
EARLIER APPLICATION NUMBER: 08/323,460
EARLIER FILING DATE: 1994-10-14
EARLIER APPLICATION NUMBER: 08/049,254
EARLIER FILING DATE: 1993-05-15
EARLIER APPLICATION NUMBER: 08/410,602
EARLIER FILING DATE: 1995-04-24
EARLIER APPLICATION NUMBER: 08/472,934
EARLIER FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 1593
TYPE: PRT
ORGANISM: Mus musculus

US-08-628-829-4

Query Match 6.8%; Score 120.5; DB 4; Length 1593;

Best Local Similarity 22.2%; Pred. No. 0.039; Matches 76; Conservative 36; Mismatches 131; Indels 99; Gaps 12;

QY 5 GATSTAAAAAARRRRRRRRRRRRRRRAVAKIYTGIRAGQDYNLPKCD 64
DB 239 GARSPPAGAPPSAA-----PSGRMEMNK-----ETLKLHMKED--REERN 279
QY 65 NNEVLKALVEAGVVEEDGTYRKQ--CKPLGEIAGTSRVTYSSQNSQPLSARQ 121
DB 280 IREKTKATCMFA-WKEHMLERRRRRGPVVKPIP--IKGDSEVNMLAAPQEGQAGSA 336
QY 122 SPIPSVQVSSSPSPSPSGRGNMMSSTFFPLPPLNGCIPSLPLRLSNSCPV----- 176
DB 337 AAPPKGRSPSPSSSPSGRSVXPS-----PQVRRRRVSPVFPQSG 377
QY 177 --TPVSSPTSKPKPL-PNWSIAKQSMALAKQSMASFPYAVSADAPFTHR----- 228
DB 378 RTTPRRAPSPGFSYSPSTSRVKNKMARLYLLOQIGNSFLIGDSDPNKTRVFI 437
QY 229 -----HOFHTLATIPEDSD-----SSTVDS 250
DB 438 GPONSCGRCACIFHLFVMLRVFQLEPBDPMLRKTLKNFEVSLFOKYHSRRSRKA 497
QY 251 GHWISFOK-----AQOQPSASWPTSPFNLYKPAPOOMSP 288
DB 498 PSRNTIOKRVSRMSNSHTLSSSTSTSSSENSIKDEEQMCP 539

RESULT 14

US-08-040-548-4

Sequence 4, Application US/08040548

Patent No. 5763209

GENERAL INFORMATION:

APPLICANT: Sukhame, Vikas P.

TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE

NUMBER OF SEQUENCES: 67

FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS

CORRESPONDENCE ADDRESS:

ADDRESS: Arnold, White & Durkee

STREET: 321 No. 5763209th Clark Street, Suite 800

CITY: Chicago

STATE: Illinois

COUNTRY: U.S.A.

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/040,548

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Coughlin, Daniel F.

REGISTRATION NUMBER: 36,111

REFERENCE/DOCKET NUMBER: arcd067

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 744-0090

TELEFAX: (312) 245-4961

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 114 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-040-548-4

Query Match 6.6%; Score 116; DB 1; Length 114;

Best Local Similarity 31.5%; Pred. No. 0.0022; Matches 29; Conservative 22; Mismatches 27; Indels 14; Gaps 3;

QY 102 SSRVTYSSQNSPLSAFQSPPIPSYQVSSSPSPSR-GEPNMMSSTFFPLRNGC 160
DB 19 SSYSPVATSYSPATTSFPSPVPSYSPGSSSTYSPHSGFSPSVATTF-----AS 72
QY 161 IP-----SLPLRLSNSCPVTTPVSSPTS 185
DB 73 VPPAFPTQVSSFPSSAGVSSSFSTGTGLSDMTA 104

RESULT 15

US-08-466-344-4

Sequence 4, Application US/08466344

Patent No. 5773583

GENERAL INFORMATION:

APPLICANT: Sukhame, Vikas P.

TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE

NUMBER OF SEQUENCES: 67

FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS

CORRESPONDENCE ADDRESS:

ADDRESS: Arnold, White & Durkee

STREET: 321 No. 5773583th Clark Street, Suite 800

CITY: Chicago

STATE: Illinois

COUNTRY: U.S.A.

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,344

FILING DATE: 06-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/040,548

FILING DATE: 31-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Coughlin, Daniel F.

REGISTRATION NUMBER: 36,111

REFERENCE/DOCKET NUMBER: arcd067

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 744-0090

TELEFAX: (312) 245-4961

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 114 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-466-344-4

Query Match 6.6%; Score 116; DB 1; Length 114;

Best Local Similarity 31.5%; Pred. No. 0.0022; Matches 29; Conservative 22; Mismatches 27; Indels 14; Gaps 3;

QY 102 SSRVTYSSQNSPLSAFQSPPIPSYQVSSSPSPSR-GEPNMMSSTFFPLRNGC 160
DB 19 SSYSPVATSYSPATTSFPSPVPSYSPGSSSTYSPHSGFSPSVATTF-----AS 72
QY 161 IP-----SLPLRLSNSCPVTTPVSSPTS 185
DB 73 VPPAFPTQVSSFPSSAGVSSSFSTGTGLSDMTA 104

Search completed: December 19, 2003, 14:27:39

Job time : 23 secs

Db 823 TCCATGCGCATTCCTAAACATCAATGCGCTTTTAATATCTTTATGCGGTTCT 882
Qy 221 AAlProAlaSerProthrhHisArgH1egInPheH1sthrLeuAlaThr1leProGluCy 240
Db 883 GCACCTGCTAGTCGACACATCGCCACAGTTTCATACCCCGCTACTATACCTGATGT 942
Qy 241 AepGluSerAepSerSerThrValAaPserGlyH1sthr1leSerPheGlnLysPheAla 260
Db 943 GATGAGTCGACCTTCCACTGCTGATTCCTGCTCATTTGATGATCTTTCAGAAgTTTGA 1002
Qy 261 GInGInGInProPheSerAlaSerMetValProThrSerProThrPheAenLeuValLys 280
Db 1003 CAACAACACCGCATTCCTCTCTATGAGCCCACTCTCTCACTTCATCTTGGA 1062
Qy 281 ProAlaProGInGInMetSerSerProAenThrAlaPheGInGInu1legLysInserSer 300
Db 1063 CCGGCGCTCAGACAGATGCTCCAAATACCTGCTTCCAGAGATGCTCAAACTCT 1122
Qy 301 GluPheLysPheGluAenSerGlnValLysProTrrPGLuG1yGluArg1leH1saPva1 320
Db 1123 GAGTTTAATTGAGATGACCAAGTTAAACCTGGAGAGAGAGATACATGATGT 1182
Qy 321 GluMetGluAepLeuGlnLeuThrlleuG1yAeng1yLysAlaArgG1y 336
Db 1183 GGTATGAGAGATCTTGAGCTTACACTTGGAATGGGAAGGCTCGTGT 1230

RESULT 3
US-09-995-938A-2
Sequence 2, Application US/0995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
TITLE OF INVENTION: HORMONE ACTION IN PLANTS
FILE REFERENCE: SALKINS.046A
CURRENT APPLICATION NUMBER: US/09/995,938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 2687
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-2

Alignment Scores:
Pred. No.: 1,226-139 Length: 2687
Score: 1671.00 Matches: 335
Percent Similarity: 74.28% Conservative: 0
Best Local Similarity: 74.28% Mismatches: 1
Query Match: 94.67% Indels: 116
Gaps: 1

US-09-995-938A-7 (1-336) x US-09-995-938A-2 (1-2687)

Qy 1 MetThrSerAepG1yAlaThrSerThrSerAlaAlaAlaAlaAlaAlaAla 20
Db 1103 ATGACTTCGATGAGTACGTCACATCAGACGCTGACGCGCGCGGCGGACACG 1162
Qy 21 AAlaArgArgLysProSerTrrArgGluArgLysAenAenArgArgArgGluArgArg 40
Db 1163 GCGAGAGAGAAAGCCGTCGTGAGAGAAAGGAGAAATTCGAGAGAGAGAAAGCGAGA 1222
Qy 41 ArgAlaValAlaAlaLys1leTrrThrG1yLeuArgAlaGInG1yAepDyrAenLeuPro 60
Db 1223 AGACTTCAGCTCGCAAGATACACTCGGCTTAAGCTCAAGGATTAATTAATTTGCT 1282
Qy 61 LysH1saCyAaPaaAenGlnValLeuLysAlaLeuCyValGluAlaG1yTrrValAla 80
Db 1283 AAACATGTCGATTAATGATGAGTCTTAAAGCTTTTGTGTAAGCTGCTTGCTTT 1342
Qy 81 GluLeuThrlleuG1yThrThrLysArgLys----- 89

Db 1343 GAAGAAGATGATCTAATATGCAA-GGTGAAGCTTCTCCATTTTTCAGANTGTA 1401
Qy 89 ----- 89
Db 1402 GCTTGTTTATTAATGANTTTTGTGATGTTGAATCGAATTCGTTGATTCATTGTGCTTA 1461
Qy 89 ----- 89
Db 1462 AATGGGTTGAATCTGAGAAATTTGAGGGTTTCTCAAGAGAAATTTGAATCATCAGAAAC 1521
Qy 89 ----- 89
Db 1522 TATGATGATCTGATTTCTCAAGTAATTAATGCGTTTCTTAAATTTAGAGTTA 1581
Qy 89 ----- 89
Db 1582 TTATGTGATGCTAAAGCTTAATCTTTATGATATATCTTGTCMAAGTCATTGCAT 1641
Qy 89 ----- 89
Db 1642 TGTGTTTCTTTGCTTACCTGATGATGATGATGTTGATGTTGTTGCTTTTG 1701
Qy 90 -----G1yCyAlaYserProLeuProG1yGln1leAlaG1yThrSerSerArgVal 105
Db 1702 TTGAGATACAGAGATGAGAGCTTACCTGCTGAGATAGCTGGGACTTCATCTCAGTA 1761
Qy 106 ThrProYrSerSerG1yAaenGInSerProLeuSerSerAlaPheGInSerPro1lePro 125
Db 1762 ACTCATTTTATCATCAGAACAGAGCCCTCTTCTACAGCTTTCMAAGTCCATCCCA 1821
Qy 126 SerTrrGlnValSerProSerSerSerSerPheProSerProSerArgG1yGluProAen 145
Db 1822 TCTTACCAAGTAAAGCCGCTTCTTCATCATTCGCCAGTCCTTCGCGGTGAACCAAT 1881
Qy 146 AaenAenMetSerSerThrPhePheProPheLeuArgAeng1yG1yleProSerSerLeu 165
Db 1882 AAACAATGTCCTCTACATTCCTTCCCTCCCAAAATAGTGAGCATTCCTCTCTCT 1941
Qy 166 ProSerLeuArg1leSerAaenSerCyProVal1ThrProProValSerSerProThrSer 185
Db 1942 CTTTCTCCAGATTCAAACGTTGCTCAGTACCCACCGCTCCTACGCGCATCTTCT 2001
Qy 186 LysAaenProLysProLeuProAenTrrPGLuSer1leAlaLysG1ySerMetAla1leAla 205
Db 2002 AAGAACCAGAAACGCTTGCCTTAACCTGAGATTAATTCGTAACCAATCCATGCCATTGCT 2061
Qy 206 LysGInSerMetValSerPheAenTrrProPheTrrAlaValSerAlaProAlaSerPro 225
Db 2062 AAACATCAATGCGGCTTTTAATATATCTTTCATAGCGGTTTCGACCTGCTACTCGG 2121
Qy 226 ThrH1saArgH1egInPheH1sthrLeuAlaThr1leProGluCyAaPGLuSerAepSer 245
Db 2122 ACACATGCGCACAGTTTCATACCTGCTACTAATCTGATGATGATGATGATGATGATGATGAT 2181
Qy 246 SerThrValAaPserG1yH1sthr1leSerPheG1yLysPheAlaG1yGlnGInProPhe 265
Db 2182 TCCACTGTTGATTTCTGCTCATTTGATGATACCTTTCGAAgTTTTCACACACACAGCATTC 2241
Qy 266 SerAlaSerMetValProThrSerProThrPheAenLeuValLysProAlaProGInGIn 285
Db 2242 TCTGCTCTATGAGGCCAACCTCTCTCACTTCAATCTTGGAACCTGCGGCTCAGCAG 2301
Qy 286 MetSerProAenThrAlaPheGInGInu1legLysInserSerGluPheLysPheGlu 305
Db 2302 ATGTCCTCAATTAATGCTGCTTCCAAAGATGCTCAAGATGCTGAGATTTAAATTTGAG 2361
Qy 306 AaenSerGlnValLysProTrrPGLuG1yGluArg1leH1saPva1MetGluAenLeu 325
Db 2362 AATAGCCAAAGTTAAACCTGGAGAGAGAGATACATGATGATGATGATGATGATGATGAT 2421
Qy 326 GluLeuThrlleuG1yAeng1yLysAlaArgG1y 336

Db 2422 GAGCTTACACTTGGAAATGGGAGGCTGCTGT 2454

RESULT 4
US-09-995-938a-1
Sequence 1, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS.046A
CURRENT APPLICATION NUMBER: US/09/995,938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2687
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-1

Alignment Scores:
Pred. No.: 5,15e-139 Length: 2687
Score: 1664.00 Matches: 334
Percent Similarity: 74.06% Conservativeness: 0
Best Local Similarity: 74.06% Mismatches: 2
Query Match: 94.28% Indels: 116
Gaps: 1

US-09-995-938a-7 (1-336) x US-09-995-938a-1 (1-2687)

Qy 1 MetThrsEaRgGlyAlaThrSerThrsEaRlAaAlaAlaAlaAlaAlaAla 20
Db 1103 ATGACTTCGATGAGTGAAGTCACTGACATCAGCAGCTGCGCGCGGACAGCAG 1162

Qy 21 ALaAArgLgLyPProSeTTPArgLgLuArgLuanAaAaArgArgArgLgArgArg 40
Db 1163 GCGAGGAGGAGCCCTCGTGGAGAAAGGAGAAATTAATCGAGAGAGAAAGAGCGAGA 1222

Qy 41 ArgAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60
Db 1223 AGAGCTGAGCTGCGCAAGATATACCTGGCTTAAGCTCAAGGATTAATTAATTTGCT 1282

Qy 61 LysHisCyAAsPaAaAaAaGluValLeuLysAlaLeuCyValGluAlaGlyTTPValVal 80
Db 1283 AAACTTTGATTAATTAATGAAGCTCTTAAGCTCTTTGTTGAAGCTGGTGGTTGTT 1342

Qy 81 GluGluAaRgLyThrThrThrArgLys----- 89
Db 1343 GAAGAAAGATGGTACTTATTCGAA--GGTGAAGACTTTCATTTTTCAGATCTGA 1401

Qy 89 ----- 89
Db 1402 GCTTGTATTATGATGTTTGTGATGTTGAATCTGAATCGTGAATTCATTTGGTTA 1461

Qy 89 ----- 89
Db 1462 AATGGCTTGAATCTGAAGATTGAAGGTTTCTCAAGTGAATTTGAATCATCAGAAAC 1521

Qy 89 ----- 89
Db 1522 TATGATGATCTGATTTCTCAAGTGAATTTATGGGTTTCTTCTAATTTAGATTA 1581

Qy 89 ----- 89
Db 1582 TTATGTGATGCTAAAGCTTAATCTTTATGTATGATGATCTGTCGAAGCTATTCAT 1641

Qy 89 ----- 89
Db 1642 TGTGTTCTTTTCTTACCTGATGATTTGATTTGATTTGTTGTTGTTGTTGTTG 1701

Qy 90 -----GlyCyAaLysPProLeuProGlyLuiLeaGlyThrsEaRgVal 105

Db 1702 TTGGAGTATCAGGATCAAGCTTTACTGCTGATAGTACGTGGACTTCATCTCGAGTA 1761

Qy 106 ThrProTySerSerGlnAaGlnSerProLeuSerSerAlaPheGlnSerProLeuPro 125
Db 1762 ACTCCATATTCATCAGAAACCAAGACCTCTTATCATCAGCTTTCAAAAGTCCATCCCA 1821

Qy 126 SerTyGlnValSerProSerSerSerSerPheProSerProSerArgGlyGluProAa 145
Db 1822 TCTTACCAAGTTACCCGCTCTTTCATCATTTCCAGAGTCTTTCTGGGGTGAACCAAT 1881

Qy 146 AsnAaMetSerSerThrPhePheProPheLeuAaGlnGlyLysIleProSerSerLeu 165
Db 1882 AACCAATGCTCTTACATTTCTTCCCTTCTCAGAAATGATGATGATCTTCTCTCT 1941

Qy 166 ProSerLeuArgLysEaAaSerCyAaProValThrProProValSerSerProThrsEa 185
Db 1942 CTTACCTCAGAAATCTCAACAGTTGTCAGTTACCCACCGGTCTCATCGCGCACTTCT 2001

Qy 186 LysAaAaProLysPProLeuProAaAaThrLysLysLysLysLysLysLysLysLys 205
Db 2002 AAGAACCCGAACCGTTCCTTACCTGAGATCTAATCGTAAAGCAATCCATGCGCATTC 2061

Qy 206 LysGlnSerMetAlaSerPheAaAaThrProPheThrAlaValSerAlaProAlaSerPro 225
Db 2062 AAACCAATCAATGGCGCTTTTAATTAATCTTTCTAATGCGGTTTCTGCACTGCTAGTCCG 2121

Qy 226 ThrHisArgHisGlnPheHisThrLeuAlaThrLysProGlyCyAaSpGlySerAaSer 245
Db 2122 AACATGCGCCAGCTTTCATACCCCGCTACTAATCTGAATGATGATGATGATGATGAT 2181

Qy 246 SerThrValAaPProSerGlyHisTTPLeuSerPheGlnLysPheAlaGlnGlnProPhe 265
Db 2182 TCCACTGTTGATTTGTCGATATGATTAAGCTTTCAAGAGTTGCAACCAACAGCCATTC 2241

Qy 266 SerLysSerMetValProThrSerProThrPheAaAaLysValLysProAlaProGlnGln 285
Db 2242 TCTGCTCTATGATGCGCAACTCTCTCTACTTCAATCTTGTGAACCTCGCGCTCAGAG 2301

Qy 286 MetSerProAaAaThrAlaAlaPheGlnLysLysGlnSerSerGluPheLysPheGlu 305
Db 2302 ATGCTCCAAATATCGCTGCTTCCCAAGATGATGATCAAGCTTCAGATTAATTTGAG 2361

Qy 306 AaAaSerGlnValLysPProTTPGluGlyLysLysLysLysLysLysLysLysLysLys 325
Db 2362 AATAGCCAAATTAACCTCGGAAAGAGAGAGATGATGATGATGATGATGATGAT 2421

Qy 326 GluLeuThrLeuGlyAaGlyLysValAaArgGly 336
Db 2422 GAGCTTACACTTGGAAATGGGAGGCTGCTGT 2454

RESULT 5
US-09-995-938a-12
Sequence 12, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS.046A
CURRENT APPLICATION NUMBER: US/09/995,938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12
LENGTH: 1213
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-12

Alignment Scores:
Pred. No.: 3,67e-127 Length: 1213

Score: 1526.50 Matches: 298
Percent Similarity: 91.67% Conservative: 10
Best Local Similarity: 88.69% Mismatches: 23
Query Match: 86.49% Indels: 5
DB: 13 Gaps: 4

US-09-995-938A-7 (1-336) X US-09-995-938A-12 (1-1213)

QY 1 MetThSerPerglylaThSerThSerAlaAlaAlaAlaAlaAlaAla 20
DB 147 ATGACGCTGACGAGCAACCTGACGCA---GCTGACCTGACGACGACGATGCG 203
QY 21 AlaArgArglyPProSerTPArgGlyArgGlyuAsnAsnArgArgArgArg 40
DB 204 ACCAGAGGAAACCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 263
QY 41 ArgAlaValAlaAlaValyleTyrThGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro 60
DB 264 AGAGCTGTCGCGGAGATTTATACGCTTACGCTCAAGTCAAGTCAATCTTCCA 323
QY 61 LysHisCysAspAsnAsnGlnValLeuLysValAlaLysValGlyTyrValVal 80
DB 324 AACATTGTGACAAACAATGAGCTTTAGGCTTTGTTCTGAAGCTGGTGGTGT 383
QY 81 GlnGluAspGlyThrThTyrArgGlyGlyCysLysProLeuProGlyGlnIleAlaGly 100
DB 384 GAGGAGAGGAACTTATTCGACAGGAGACAAAGCTTACCTGATGACATGCTGGA 443
QY 101 ThrSerSerArgValThrProTyrSerSerGlnAsnGlnSerProLeuSerSerAlaPhe 120
DB 444 TCATCTTCTCGAGAACTCTTACTCTCCATACCAAGTCTCTTCTTCTTCACTTT 503
QY 121 GlnSerProIleProSerTyrGlnValSerProSerSerSerPheProSerProSer 140
DB 504 GATAGCCCCATTTATCTTACCAAGTCACTCTTCTTCTTCAATCCGAGCTTCT 563
QY 141 Arg---GlyGluProAsnAsnAsnMetSerSerThrPhePheProLeuAlaGly 159
DB 564 CGAGTTGGATGACCAAT-----ATCTCCAAATCTTCTTCTTCTTCTTCTT 617
QY 160 GlyIleProSerSerLeuProSerLeuArgIleSerAsnSerCysProValThrProPro 179
DB 618 GGTATCTCTTCATCGCTCTCTCCATTAATCTCAACAGTCTCTGCTGCTCACTCCA 677
QY 180 ValSerSerProTherLysAsnProLysProLeuProAsnThrPgluSerIleAlaLys 199
DB 678 GTGTCACTCCCACTTCTAGAAACCCCAACCAATGCTTCTTCTTCTTCTTCTT 737
QY 200 GlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPheAsnTyrProPheTyrAla 218
DB 738 CAATCATCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 797
QY 219 ValSerAlaProLysSerProThrThLysGlnGlnPheHisThrLeuAlaThrIlePro 238
DB 798 GCTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
QY 239 GlnCysAspGlnSerAspSerSerThrValAspSerGlyHisThrIleSerPheGlnLys 258
DB 858 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 917
QY 259 PheAlaGlnGlnGlnProPheSerAlaSerMetValProTherSerProThrPheAsnLeu 278
DB 918 TTTCACAACAACAACGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977
QY 279 ValLysProAlaProGlnInMetSerProAsnThrAlaAlaPheGlnGlnIleGlyGln 298
DB 978 GCGAAACCTGCAACGACAAATGCTTCCAAACAGCAAGCAATCCAAAGATGATGCA 1037
QY 299 SerSerGlnPheLysPheGlnAsnSerGlnValLysProTyrGlnGlyGlnArgIleHis 318
DB 1038 ACCTCCGACTTAAAGTTTGAAGAACCCCAAGTAAACCTGGAAGGAGAGATCCAT 1097
QY 319 AspValGlyMetClnAspLeuGlnLeuThrLeuGlyAsnGlyLysAla 334

DB 1098 GATGCTGCTATGAGATCTAGAGCTCAGCTTGAAATGTAAGCT 1145

RESULT 6

US-09-995-938A-11
Sequence 11, Application US/0995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOHANN CHOR
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS, 046A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1213

ALIGNMENT SCORES:
Pred. No.: 1,556-126 Length: 1213
Score: 1519.50 Matches: 297
Percent Similarity: 91.37% Conservative: 10
Best Local Similarity: 88.39% Mismatches: 24
Query Match: 86.09% Indels: 5
DB: 13 Gaps: 4

US-09-995-938A-7 (1-336) X US-09-995-938A-11 (1-1213)

QY 1 MetThSerPerglylaThSerThSerAlaAlaAlaAlaAlaAlaAla 20
DB 147 ATGACGCTGACGAGCAACCTGACGCA---GCTGACCTGACGACGACGATGCG 203
QY 21 AlaArgArglyPProSerTPArgGlyArgGlyuAsnAsnArgArgArgArg 40
DB 204 ACCAGAGGAAACCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 263
QY 41 ArgAlaValAlaAlaValyleTyrThGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro 60
DB 264 AGAGCTGTCGCGGAGATTTATACGCTTACGCTCAAGTCAAGTCAATCTTCCA 323
QY 61 LysHisCysAspAsnAsnGlnValLeuLysValAlaLysValGlyTyrValVal 80
DB 324 AACATTGTGACAAACAATGAGCTTTAGGCTTTGTTCTGAAGCTGGTGGTGT 383
QY 81 GlnGluAspGlyThrThTyrArgGlyGlyCysLysProLeuProGlyGlnIleAlaGly 100
DB 384 GAGGAGAGGAACTTATTCGACAGGAGACAAAGCTTACCTGATGACATGCTGGA 443
QY 101 ThrSerSerArgValThrProTyrSerSerGlnAsnGlnSerProLeuSerSerAlaPhe 120
DB 444 TCATCTTCTCGAGAACTCTTACTCTCCATTAACCAAGTCTCTTCTTCTTCACTTT 503
QY 121 GlnSerProIleProSerTyrGlnValSerProSerSerSerPheProSerProSer 140
DB 504 GATAGCCCCATTTATCTTACCAAGTCACTCTTCTTCTTCAATCCGAGCTTCT 563
QY 141 Arg---GlyGluProAsnAsnAsnMetSerSerThrPhePheProLeuAlaGly 159
DB 564 CGAGTTGGATGACCAAT-----ATCTCCAAATCTTCTTCTTCTTCTTCTT 617
QY 160 GlyIleProSerSerLeuProSerLeuArgIleSerAsnSerCysProValThrProPro 179
DB 618 GGTATCTCTTCATCGCTCTCTCCATTAATCTCAACAGTCTCTGCTGCTCACTCCA 677
QY 180 ValSerSerProTherLysAsnProLysProLeuProAsnThrPgluSerIleAlaLys 199
DB 678 GTGTCACTCCCACTTCTAGAAACCCCAACCAATGCTTCTTCTTCTTCTTCTT 737

```

QY 200 GlnSerMetAlaIle---AlaValGlnSerMetAlaSerPheAsnTyrProPheTyrAla 218
Db 738 CAATTCATGCTCATGCTGCTGCTAAACAGTCAATGACTCTTTTGAACCTACCCGTTTATGCG 797
QY 219 ValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrIlePro 238
Db 798 GTGCTGCACTGCGCAAGCTCTACTCATCATGCGCAAGTTCAGTCCGCGCTACTACT 857
QY 239 GluCyAspGluSerAspSerSerThrValAspSerGlyHisTyrIleSerPheGlnLys 258
Db 858 GAATGTGATGAGTCTGACTCTTCCACTGTGATTCGTGCTCATGGATTAAGCTTTCAAAAG 917
QY 259 PheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsnLeu 278
Db 918 TTTCACACACACACACCACTTCTCTCTCTATGTCGCCAACCTCGCTACCTTCAATCTC 977
QY 279 ValLysProAlaProGlnGlnMetSerProAsnThrAlaAlaPheGlnGlnIleGlyGln 298
Db 978 GTGAAACCTGCAACCAACCAATGCTCTCCAAACACAGCAATCCAAAGATTTGGTCA 1037
QY 299 SerSerGluPheLysPheGluAsnSerGlnValLysProTyrGluGlyLysArgIleHis 318
Db 1038 AGCTCCGAGTTTAAGTTTGAAGAACAGCCAAATTAAAGCCATGGAGGAGGAGATCCAT 1097
QY 319 AspValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGlyLysAla 334
Db 1098 GATGTGCTATGAGAGATCTAAGCTTCAAGCTTGAAATGTAAGCT 1145

```

RESULT 7

```

US-09-995-938A-9
: Sequence 9, Application US/0995938A
: Publication No. US20030150026A1
: GENERAL INFORMATION:
: APPLICANT: JOANNE CHORY
: APPLICANT: ZHIYONG WANG
: TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
: TITLE OF INVENTION: HORMONE ACTION IN PLANTS
: FILE REFERENCE: SALKINS.046A
: CURRENT APPLICATION NUMBER: US/09/995,938A
: CURRENT FILING DATE: 2001-11-27
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 3000
: TYPE: DNA
: ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-9

```

Alignment Scores:

```

Pred. No.: 7,62e-119 Length: 3000
Score: 1439.00 Matches: 298
Percent Similarity: 69.68% Conservative: 10
Best Local Similarity: 67.42% Mismatches: 23
Query Match: 81.53% Indels: 112
DB: 13 Gaps: 6

```

US-09-995-938A-7 (1-336) x US-09-995-938A-9 (1-3000)

```

QY 1 MetThrSerAspGlyValaThrSerThrSerAlaAlaAlaAlaAlaAlaAlaAla 20
Db 1337 ATGAGCTCTGACGGAGCAACGTCGACGTA--GCTGCAGCTGCACAGCAGCGATGGCG 1393
QY 21 AlaAspArgLysProSerTyrPArgLysArgLysAsnArgArgArgLysArgArg 40
Db 1394 ACGAGAGAGAAACCGTCGTGAGAGAGAGAGAGAACATCTGAGGAAGAGCGCGAGGA 1453
QY 41 ArgAlaValAlaAlaLysIleTyrThrGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro 60
Db 1454 AAGAGCTGTGCGCGCAAGATTTTACTGCTTGAAGCTCAAGTCACTACATCTTTCCA 1513
QY 61 LysHisCysAspAsnAsnGluValLleuLysAlaLeuCysValGluAlaGlyTyrValVal 80
Db 1514 AAACATTTGACAAACAATGAGGTTCTTAAGGCTTTTGTCTGAACCTGGTTGGTTGT 1573

```

```

QY 81 GluGluAspGlyThrThrTyrArgLysGly----- 90
Db 1574 GAAGAAAGCGGAACACTTATTCGCA--GATCAGTATCAAAAGCATTTTACTTAGATCTG 1632
QY 90 ----- 90
Db 1633 ATGTTATCTGATGATTTAGCTGTTGATCTGAAGATTGGATTGAAATTGGCAATTG 1692
QY 90 ----- 90
Db 1693 GGATTTCTGGCTATGAATTCAGAGTTTCTAGTGAAGAGTCTAGTTTATTCTTAAA 1752
QY 90 ----- 90
Db 1753 TTGATCGAGATCTCTCGGAGAAAGTACCTTGAAGGTTCTTACTAATTGAGAA 1812
QY 90 ----- 90
Db 1813 CCGAATTAGCTTTACTTCACTTGGTACTATATTTAGATCTCTCTTAGCTTTGATT 1872
QY 91 ---Cys-----LysProLeu 94
Db 1873 GATGTGACATGTGATGTTTTTGGATTTGTTCTATGACAAACAGGACAAAGCCTCTA 1932
QY 95 ProGlyGluIleAlaGlyThrSerSerArgValThrProTyrSerSerGlnAsnGlnSer 114
Db 1933 CCTGTGACATGGCGATCACTTCTCGAGCAATCTTACTCTTCCATTAACCAAGT 1992
QY 115 ProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerProSerSerSer 134
Db 1993 CCTCTTTCTTCCACTTTGATAGCCCATCTTATCTTACCAAGTCACTCTTCTTCT 2052
QY 135 SerPheProSerProSerArg---GlyLysProAsnAsnAsnMetSerThrPhePhe 153
Db 2053 TCATTCGAGAGCTTCTCGAGTGTGATTCACACAAAT-----ATCTCCAAATCTTC 2106
QY 154 ProPheLeuArgAsnGlyGlyLysProSerSerLeuProSerLeuArgIleSerAsnSer 173
Db 2107 CCTTTCTCCAGAAATGGTGTGTTCTTCATGCTTCCCTCCACTTGAATTCGAAACAGT 2166
QY 174 CysProValThrProProValSerSerProThrSerLysAsnProLysProLeuProAsn 193
Db 2167 GCTCGTCTACCTCCACAGTGTCAATCCCACTTCTTGAAGCCCAACCATTTGCTACT 2226
QY 194 TrpGluSerIleAlaLysGlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPhe 212
Db 2227 TGGGAATCTTTTACCAACCAATCCATGTCATGGCTGTAACAGTCAATGACTTCTTGG 2286
QY 213 AsnTyrProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHis 232
Db 2287 AACTACCCCTTTTATGCGGTGTCTGACCTGCCAATCTTCACTCATATCGCCAGTTCCAT 2346
QY 233 ThrLeuAlaThrIleProGluCysAspGluSerAspSerSerThrValAspSerGlyHis 252
Db 2347 GCTCTGCTACATACCTGAATGTAGTGTGACTTTCACCTGTTGATTTCTGGTCAAT 2406
QY 253 TrpLysSerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSerMetValProThr 272
Db 2407 TGGATAGCTTCAAAAGTTTGCACAAACAGCCATTTCTGCTCTATGAGTGCACAC 2466
QY 273 SerProThrPheAsnLeuValLysProAlaProGlnGlnMetSerProAsnThrAlaAla 292
Db 2467 TCGCTTACCTTCAATCTCGTGAATCTGCACCAACCAATTTCTCCAAACACAGCGCA 2526
QY 293 PheGlnGluIleGlyGlnSerSerGluPheLysPheGluAsnSerGlnValLysProTyr 312
Db 2527 ATCCAAAGATTTGTAAGGCTCCGAGTTTAACTTTGAGAACGCCAAAGTTAAAGCCATGG 2586
QY 313 GluGlyGluArgIleHisAspValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGly 332
Db 2587 GAAGGAGAGAGATTCATGATGTGCTATGAGAGATCTAGAGCTCAAGCTTGGAAATGATGT 2646

```

QY 333 LysAla 334
 Db 2647 AAAGCT 2652

RESULT 8

US-09-995-938a-3
 ; Sequence 3, Application US/09995938A
 ; Publication No. US20030150026A1
 ; GENERAL INFORMATION:
 ; APPLICANT: JOANNE CHORY
 ; APPLICANT: ZHIYONG WANG
 ; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
 ; TITLE OF INVENTION: HORMONE ACTION IN PLANTS
 ; FILE REFERENCE: SALKINS 046A
 ; CURRENT APPLICATION NUMBER: US/09/995, 938A
 ; CURRENT FILING DATE: 2001-11-27
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PaeSeQ for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 3000
 ; TYPE: DNA
 ; ORGANISM: ARABIDOPSIS THALIANA
 ; US-09-995-938a-3

Alignment Scores:

Pred. No.: 3,22e-118 Length: 3000
 Score: 1432.00 Matches: 297
 Percent Similarity: 69.46% Conservative: 10
 Best Local Similarity: 67.19% Mismatches: 24
 Query Match: 81.13% Indels: 112
 DB: 13 Gaps: 6

US-09-995-938a-7 (1-336) x US-09-995-938a-3 (1-3000)

QY 1 MetThrSerAspAlaValThrSerThrSerAlaAlaAlaAlaAlaAlaAla 20
 Db 1337 ATGACGCTGACCGAGCAAGCTGACGCTCA---GCTCAGCTGACGACGACGATGCG 1393
 QY 21 AlaArgArgLysProSerTrpArgGluArgGluAsnArgArgArgArgArg 40
 Db 1394 ACCAGAGAGAAACCGCTGCTGAGAGAGAGAGAAACATCGAGAAAGACGCGGAGA 1453
 QY 41 ArgAlaValAlaAlaValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 Db 1454 AGAGCGGTGGCGGAGAAATTAATGCTTGAAGCTTCTTGAAGCTTCTTGAAGCTT 1513
 QY 61 LysHisCysAspAsnAsnGluValLeuLeuValAlaLeuCysValGluAlaGlyTrpVal 80
 Db 1514 AAACATTTGTACAAACAATGAGTCTTAAAGCTCTTTGTTCTGAAAGCTGGTGGTGT 1573
 QY 81 GluGluAspGlyThrThrTrpArgLysGly----- 90
 Db 1574 GAAGAGACGAGAACTACTATCCAA-GATCAGATCAAAAGCATTTTACTAGATCTG 1632
 QY 90 ----- 90
 Db 1633 ATGTATCTGATGATTAGCTGTGATCTGAAGATTGATTGAATTGTCAAATG 1692
 QY 90 ----- 90
 Db 1693 GGATTTCTTGCTATGAATTCGAGGTTTACTGAGAAAGCTCAAGTTTATTCTAAA 1752
 QY 90 ----- 90
 Db 1753 TTGGATCGAGATCTTGGCGGAGAAAGTACCTTTAGGCTTCTTACTAATTGAGAA 1812
 QY 90 ----- 90
 Db 1813 CGGAATTAGCTTACTTCACTGTGTTACTATATTAGATCTCTCTTACTGTTGATT 1872
 QY 91 ---Cys-----LysProLeu 94
 Db 1873 GATTGACATGTGTGATGTTTGGTATTGTTCTATGAGCAACAGGACACAGGCTCTA 1932

QY 95 ProGluIleAlaGlyThrSerSerArgValThrProTyrSerSerGlnAsnGlnSer 114
 Db 1933 CCGTGATACATGGCTGGATCATCTTCTGACCAATCTCTTACTCTTCCATACCAAGT 1992
 QY 115 ProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerProSerSer 134
 Db 1993 CCTCTTCTTCCACTTTGATGAGCCCATCTTATTTACCAAGTCAGTCTTCTTCTTCT 2052
 QY 135 SerPheProSerProSerArg---GlyIleProAsnAsnAsnMetSerSerThrPhe 153
 Db 2053 TCATTCGCACTCTTCTTCCAGTGTGATCACAACAT-----ATTCACAACATCTTC 2106
 QY 154 ProPheLeuArgAsnGlyIleProSerSerLysAsnProSerLeuArgIleSerAsnSer 173
 Db 2107 CCTTCTCAGGAATGTGATTCCTTCAATCGCTTCTTCCACTTGAATCAACAACGT 2166
 QY 174 CysProValThrProProValSerSerProThrSerLysAsnProLysProLeuProAsn 193
 Db 2167 GCTCTGTCACTCCACAGTGTCAATCCCACTTTAGAAACCCCAACCAATGCTTACT 2226
 QY 194 TrpGluSerIleAlaLysGlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPhe 212
 Db 2227 TGGGAATCTTTTACCAACAATCATGCTGCTGCTAACAAGTCAATGATCTTCTTGG 2286
 QY 213 AsnTyrProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHis 232
 Db 2287 AACTACCCGTTTATGCGGTGTGTGCACCTGCCAGTCTTACTATCATTCGCGAGTTCAT 2346
 QY 233 ThrLeuAlaThrIleProGluCysAspGluSerAspSerSerThrValAspSerGlyHis 252
 Db 2347 GCTCCGGCTACTATCACTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2406
 QY 253 TrpIleSerPheGlnLysPheAlaGlnGlnGlnPheSerAlaSerMetValProThr 272
 Db 2407 TGGATAGCTTTCAAAAGTTTGACAAACAAGCATTTCTTCTGCTTATGTGCTCAACC 2466
 QY 273 SerProThrPheAsnLeuValLysProAlaProGlnIleMetSerProAsnThrAlaAla 292
 Db 2467 TCGGATACCTTCAATCTCGTGAACCTGACACACAGCAATTTGTTCCAAACACGACCA 2526
 QY 293 PheGlnIleIleGlyGlnSerSerGluPheLysPheGluAsnSerGlnValLysProTrp 312
 Db 2527 ATCCAGAGATTTGCTCAAGCTCCGAGTTTAAGTTTGAAGAACGCCAAGTTAAGCATG 2586
 QY 313 GluGlyGluAspGlyLeuHisAspValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGly 332
 Db 2587 GAAGGGAGAGATCCATGATGTGGCTATGAGAGATCTAGAGCTCAGCGCTTGAAATGCT 2646
 QY 333 LysAla 334
 Db 2647 AAAGCT 2652

RESULT 9

US-09-938-842a-385
 ; Sequence 385, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Krepes, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227, 866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264, 647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300, 111
 ; PRIOR FILING DATE: 2001-06-22


```

: FEATURE:
: NAME//KEY: CDS
: LOCATION: (201) .. (968):
US-10-017-161-2179

Alignment Scores:
Pred. No.: 0.000164 Length: 1168
Score: 154.00 Matches: 62
Percent Similarity: 43.14% Conservative: 26
Best Local Similarity: 30.39% Mismatches: 95
Query Match: 8.73% Indels: 21
DB: 13 Gaps: 4

US-09-995-938A-7 (1-336) x US-10-017-161-2179 (1-1168)

QY 92 LysProLeuProGlyGluIleAlaGlyThrSerSerArgValThrProTyrSerSerGln 111
DB 750 AAACCATCACCACCA-----ACACCATCATCATCATCATCATCATCATCATCATCA 706

QY 112 AsnGlnSerProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerPro 131
DB 705 CCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 646

QY 132 SerSerSerSerPheProSerProSerArgGlyGluProAsnAsnAsnMetSerSerThr 151
DB 645 TCACGAGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCTCA 566

QY 152 PhePheProPheLeuArgAsnGlyGlyLeProSerSerLeuProSerLeuArgIleSer 171
DB 585 CCACCACCATCATCATCCACCATCATCATCATCATCATCATCATCATCATCATCATCATCA 526

QY 172 AsnSerCysProValThrProProValSerSerProThrSerTyrAsnProLysProLeu 191
DB 525 TCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 466

QY 192 ProAsnTrpGluSerIleAlaLysGlnSerMetAlaIleAlaLysGlnSerMetAlaSer 211
DB 465 CCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 406

QY 212 PheAsnTrpProPheTyrAlaValSerAlaProAlaSerProThr-HisArgHisGlnPb 231
DB 405 TCATCATCATCATCA-----TCCTCATCATCATCATCATCATCATCATCATCATCA 352

QY 231 eHis-ThrLeuAlaThrIle-ProGluCysArgGluSerArgSerSerThrValAspSer 250
DB 351 TCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 292

QY 251 GlyHisTrpIleSerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSerMetVal 270
DB 291 TCA-----CCATGCTCTTACCACCATCATCATCATCATCATCATCA 256

QY 271 ProThrSerProThrPheAsnLeuValLysProAlaProGlnGlnMetSerProAsnThr 290
DB 255 TCATCATCATCATCATCATCA-----TCATCATCATCATCATCATCATCATCATCATCA 205

QY 291 AlaIaL 292
DB 204 TCATCA 199

RESULT 13
US-10-156-761-930
: Sequence 930 Application US/10156761
: Publication No. US20030119018A1
: GENERAL INFORMATION:
: APPLICANT: OMOURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, MASAHIRA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-262

```



```

QY 65 -----AsnAngluValIleuysalaleuCyvalGluIaGlyTrpValValGI 81
DB 1174521 CCACCCCTCCACGACGACACTCTGCACTGCTG----- 1174488
QY 81 uGluAerGlyThrThrTyrrArgLyGlyCylysProLeuPro-----GlyGluI 98
DB 1174487 -GGAGGATCACCACCCACAGCCGTCACAGATCCGCTTCTTCCACCGCTCAGGACCT 1174429
QY 98 leaIaGlyThrSerSerArgValThrPro-----TyrSerGlnAsnGlnSerProI 116
DB 1174428 GGCTGTGATACACCAACCTGAGCCGCTACTGGTACCGACACTCCACAGCCCT 1174369
QY 116 euSerSerAlaPheGlnSer-----ProIleProSerTyrglnValSerProSerSers 135
DB 1174368 CCGCTTCACCCACGCGCTGACGACCTGACGAGGACGACCGACCGCTTCATCGAAT 1174309
QY 135 erPheProSerProSerArgLyGlyProAsnAsnMetSerSerThrPhePhePro 155
DB 1174308 CAGTCCCAACCCACCTGCTGCTCCGCTGACGAGACACACCGAAGAA----- 1174259
QY 155 heLeuArgAngLyGlyIleProSerSerLeuProSerLeuArgIleSerAsnSerCyas 175
DB 1174258 -----CCACCGAAGAAATCACCAGGACCGGACCGCTCCGCGGCGCA 1174216
QY 175 rovalThrProProValSerSerProThrSerLyAsnProLyPProLeuProAsnTrpG 195
DB 1174215 CACGACACCCACCGCTTCTTCAACCGCTCCGCGGACACCCACCA----- 1174169
QY 195 luSerIleAlaLyGlnSerMetAlaIleAlaLyGlnSerMetAlaSerPheAsnTrp 215
DB 1174168 -----CCGGCATTCGACACCC 1174153
QY 215 roPheTyraIaValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeuA 235
DB 1174152 CACGACCTGACACACCTACACCAACCCACCCACCC-----CCACACACCACT 1174096
QY 235 laThrIleProGluCyAsnArgLySerSerSerThrValAspSerGlyHisTrpIles 255
DB 1174095 CGACCTGCCACCT--ACCCCTTCACACACGACGACCTACTGCTCAACAC----- 1174046
QY 255 erPheGlnLyPheAlaGlnGlnInProPheSer-----AlaSerMetValProT 272
DB 1174045 -----CCACACACACACACGACCTGACACCGGCTCAGCCACCA 1174000
QY 272 hrSerProThrPheAsnLeuValLyPProAlaPro 283
DB 1173999 CCACCCCTCCTCACCACGACACTCACCCTCGCA 1173965

RESULT 15
US-10-029-386-20619
: Sequence 20619, Application US/10029386
: Publication No. US20030194704A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
: FILE REFERENCE: AROMICA-X-2
: CURRENT APPLICATION NUMBER: US/10/029,386
: NUMBER OF SEQ ID NOS: 34288
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 20619
: LENGTH: 506
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC002070.1 SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN H129, SIGNAL = 0.79
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2

```

```

: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
US-10-029-386-20619

Alignment Scores:
Pred. No.: 0.000548 Length: 506
Score: 145.00 Matches: 59
Percent Similarity: 41.67% Conservative: 21
Best Local Similarity: 30.73% Mismatches: 81
Query Match: 8.10% Indels: 31
DB: 13 Gaps: 3

US-09-995-938a-7 (1-336) x US-10-029-386-20619 (1-506)
QY 103 SerArgValThrProTyrrSerSerGlnAsnGlnSerProLeuSerSerAlaPheGlnSer 122
DB 5 TCATCATCATCATCATCATCTTCTATCATCATCATCATCATCATCATCATCATCATCA 64
QY 123 ProIleProSerTyrglnValSerProSerSerSerPheProSerProSerArgLy 142
DB 65 CCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCTCA 124
QY 143 GluProAsnAsnMetSerSerThrPhePheProPheLeuArgAngLyIlePro 162
DB 125 CCACCATCATCATCATCATCATCTTCTCACCACCAACCATCATCATCATCATCATCA 184
QY 163 SerSerLeuProSerLeuArgIleSerAsnSerCyProValThrProProValSerSer 182
DB 185 TCATCTTCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 244
QY 183 ProThrSerLyAsnProLyPProLeuProAsnTrpGluSerIleAlaLyGlnSerMet 202
DB 245 CCATCATCTTCCACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 277
QY 203 AlaIleAlaLyGlnSerMetAlaSerPheAsnTrpProPheTyraIaValSerAlaPro 222
DB 278 -----TCCTCACACCAACCATCATCATCATCATCATCATCATCATCA 307
QY 223 AlaSerProThrHisArgHisGlnPheHis-ThrLeuAlaThrIle-ProGluCyAsp 241
DB 308 CCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 367
QY 242 GluSerAspSerSerThrValAspSerGlyHisTrpIleSerPheGlnLyPheAlaGln 261
DB 368 CCATCA---TCATCACACCATCATCATCATCATCATCATCATCATCATCATCATCA 424
QY 262 GlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsnLeuValLyPPro 281
DB 425 TCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA----- 463
QY 282 AlaProGlnGlnMetSerProAsnThrAla 291
DB 464 ---CCATCATCATCATCATCATCATCATCATCATCA 490

```

Search completed: December 20, 2003, 07:57.34
Job time : 3656 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 19, 2003, 22:30:12, Search time 4384 Seconds

(without alignments)
3135.409 Million cell updates/sec

Title: US-09-995-938A-7
Perfect score: 1765
Sequence: 1 MTSDBGATSTGTAATAAAAAA.....IHDVGMEDLEITLNGKARG 336

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US0995938/runat.19122003.142606.4237/app_query.fasta_1.519
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=1 -MATRIX=Diosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0995938 @CGN 1 1 3108 @runat.19122003.142606.4237 -NCPU=6 -ICPU=3
-NO MMAP -LARGEODRBY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb_da: *
2: gb_hg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sct: *
12: gb_sy: *
13: gb_un: *
14: gb_vt: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pac: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sct: *
28: em_un: *

29: em_vt: *
30: em_hg_hum: *
31: em_hg_inv: *
32: em_hg_other: *
33: em_hg_mus: *
34: em_hg_pln: *
35: em_hg_rnd: *
36: em_hg_mam: *
37: em_hg_vtc: *
38: em_sy: *
39: em_hggo_hum: *
40: em_hggo_mus: *
41: em_hggo_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1758	99.6	1501	8	AY087257 Arabidops
2	1755	99.4	1011	8	AY093747 Arabidops
3	1755	99.4	1476	8	AY065049 Arabidops
4	1664	94.3	2687	8	AF494338 Arabidops
5	1664	94.3	97263	8	AC013258 Arabidops
6	1519.5	86.1	1340	8	AY065041 Arabidops
7	1513.5	85.8	1257	8	AF134217 Arabidops
8	1511.5	85.6	1359	8	AY086340 Arabidops
9	1432	81.1	120977	8	AC025808 Genomic s
10	1060	60.1	1494	8	AF395901 Lycopersi
11	1035	58.6	945	8	AF372937 Arabidops
12	872.5	49.0	1389	8	AY050394 Arabidops
13	864.5	48.8	957	8	AY097357 Arabidops
14	860.5	48.5	144480	8	AP004276 Oryza sat
15	856	46.7	831	6	AX505690 Sequence
16	825	46.7	831	6	BT006310 Arabidops
17	825	46.7	1272	8	BT006310 Arabidops
18	825	46.7	1272	8	BT006310 Arabidops
19	788	44.6	84196	8	AT7385 Arabidops
20	788	44.6	89469	8	AT7183 Arabidops
21	721.5	40.9	198354	8	ATAP22 Arabidops
22	721.5	40.9	198354	8	ATAP22 Arabidops
23	642.5	36.4	411	8	AY074829 Arabidops
24	585.5	33.2	978	8	AY090331 Arabidops
25	585.5	33.2	1616	8	AY050430 Arabidops
26	561.5	31.8	1116	6	AX653341 Sequence
27	541	30.7	855	6	AX651445 Sequence
28	541	30.7	1217	8	AY088379 Arabidops
29	541	30.7	1258	8	AK118850 Arabidops
30	461.5	26.1	90425	8	F9K20 Arabidops
31	416.5	23.6	15815	8	AP003105 Oryza sat
32	339.5	19.2	192110	2	AP004070 Oryza sat
33	326.5	18.5	166126	2	AP003686 Oryza sat
34	313.5	17.8	176734	2	AP003682 Oryza sat
35	299	16.9	119111	8	ATF13C5 Arabidops
36	299	16.9	199075	8	ATCRHIV49 Arabidops
37	293.5	16.6	133330	2	OSJN00266 Arabidops
38	293.5	16.6	148951	2	OSJN00241 Arabidops
39	283.5	16.1	161277	8	AP002972 Oryza sat
40	255	14.4	374	8	AY200612 Arabidops
41	233	13.2	163095	8	AP002523 Arabidops
42	204.5	11.6	2070	8	BT006482 Arabidops
43	204.5	11.6	2272	8	AK117140 Arabidops
44	201	11.4	125008	2	AP005546 Oryza sat
45	195.5	11.1	109135	8	AP003408 Oryza sat

RESULT 1

ALIGNMENTS

LOCUS	AY087257	1501 bp	mRNA	linear	PLN 14-APR-2003
DEFINITION	Arabidopsis thaliana clone 33367 mRNA, complete sequence.				
ACCESSION	AY087257				
VERSION	AY087257.1	GI:21405981			
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	1 (bases 1 to 1501) Hase, B.J., Volkovskiy, N., Town, C.D., Troukhan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O. and Szidzberg, S.L.				
TITLE	Full-length messenger RNA sequences greatly improve genome annotation				
JOURNAL	Genome Biol. 3 (6), RESEARCH0029 (2002)				
MEDLINE	22088475				
PUBMED	12093376				
REFERENCE	2 (bases 1 to 1501) Brewer, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.				
TITLE	Full-length cDNA from Arabidopsis thaliana				
JOURNAL	Unpublished				
AUTHORS	3 (bases 1 to 1501) Brewer, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.				
TITLE	Direct Submission				
COMMENT	Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Geneset carried out the library production and sequencing of the full-length clones. Ceres, Inc, carried out the clustering of the 5' sequences, selection of clones, and sequence assembly. Location/Qualifiers 1..1501 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="33367" 187..1197 /codon_start=1 /product="unknown" /protein_id="AAM64812.1" /db_xref="GI:21592862" /translation="MTSDQATSTAAAAAARRKRSRENNRRRRRAVAVAKITVGLAAGDYNLPKHCDNNEVTKALCVAGAVVEDDCTYRGCKPLGLAGTSRKVTPYSSQNSQPLSAPFOSPIPPSVQVPSPPSSPSPSPSGRBNNMSTPFPFLANGSPVPSLPLINSCVPTPPVSPSTKSNPLPNMWSINKOSMAFKOSWAFNPPVAFVSPASPPHRRHQFHTPATIPEDCEDSDSTVDSGHWISFOKAOOOPFASVWPPSPDSNVLKVAAPQOQMSFNTAFOEIGQSESEFKFENSQVKEWEEERIDHVGMDLETLTGKXKRG"				

[illegible]

LOCUS	LOCUS	DEFINITION	DEFINITION	VERSION	VERSION	KEYWORDS	KEYWORDS	SOURCE	SOURCE	ORGANISM	ORGANISM
AY093747	AY093747	1011 bp	1011 bp	linear	linear	PLN 13-APR-2002	PLN 13-APR-2002				
Arabidopsis thaliana	Arabidopsis thaliana	Atig75080/P9510_7	Atig75080/P9510_7	mrna	mrna	complete cds.	complete cds.				
AY093747	AY093747	1	1	GI:20147314	GI:20147314						
FLI CDNA.	FLI CDNA.										
Arabidopsis thaliana	Arabidopsis thaliana	(thale cress)	(thale cress)								
Arabidopsis thaliana	Arabidopsis thaliana										
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;										
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;										
rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi	rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi										
1 (bases 1 to 1011)	1 (bases 1 to 1011)										
Shim, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,	Shim, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,										
Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,	Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,										
Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G.,	Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G.,										
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,	Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,										
Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,	Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,										
Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C.,	Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C.,										
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,	Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,										
Theologis, A. and Ecker, J.R.	Theologis, A. and Ecker, J.R.										
Arabidopsis ORF clones	Arabidopsis ORF clones										
Unpublished	Unpublished										
2 (bases 1 to 1011)	2 (bases 1 to 1011)										
Shim, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,	Shim, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,										
Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,	Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,										
Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G.,	Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G.,										
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,	Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,										
Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,	Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,										
Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C.,	Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C.,										
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,	Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,										
Theologis, A. and Ecker, J.R.	Theologis, A. and Ecker, J.R.										
Direct Submission	Direct Submission										
Submitted (23-MAR-2002) Salk Institute Genomic Analysis Laboratory	Submitted (23-MAR-2002) Salk Institute Genomic Analysis Laboratory										
(SIGMAL), Plant Biology Laboratory, The Salk Institute for	(SIGMAL), Plant Biology Laboratory, The Salk Institute for										
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,	Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,										
USA	USA										
RIKEN Genomic Sciences Center (GSC) members carried out the	RIKEN Genomic Sciences Center (GSC) members carried out the										
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN	collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN										
Arabidopsis Full-Length cDNA' : Seki, M., Narusaka, M., Ishida, J.,	Arabidopsis Full-Length cDNA' : Seki, M., Narusaka, M., Ishida, J.,										
Satou, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J.,	Satou, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J.,										
Hayashizaki, Y. and Shinozaki, K.	Hayashizaki, Y. and Shinozaki, K.										
The Salk, Stanford, PGSC (SSP) Consortium members constructed and	The Salk, Stanford, PGSC (SSP) Consortium members constructed and										
sequenced the PUN1 (ORF) clones using the RAPL cDNAs: Shim, P.,	sequenced the PUN1 (ORF) clones using the RAPL cDNAs: Shim, P.,										
Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L.,	Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L.,										

[illegible]

DB 841 CCTGCCCTCAGCAGATGCTCCAAATATCTGCTCCCTCCAGAGATTGCTCAACTCT 900

QY 301 GIupheliaepheliausensergivalylpfpctfngluygluargilehisapval 320

DB 901 GAGTTAAATTTGAGATAGCCAAAGTTAAACCTCGGAGAGAGAGATACATGATG 960

QY 321 GIymetgluapleugluileuthrluenglylaenglylaargly 336

DB 961 GGATATGAGATCTTGAAGCTTACCTTGGAAGGAGAGCTGTGT 1008

RESULT 3

AY065049 1476 bp mRNA linear PLN 07-JAN-2002

LOCUS Arabidopsis thaliana Atg1975080/P9E10_7 mRNA, complete cds.

DEFINITION AY065049.1 GI:18086460

VERSION Arabidopsis thaliana (thale cress)

KEYWORDS Arabidopsis thaliana

SOURCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1476)

AUTHORS Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Arabidopsis cDNA clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1476)

AUTHORS Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Direct Submision

JOURNAL Submitted (30-NOV-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bower, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Phan, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A. and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers

1. 1476

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="RAFL04-20-E20 (R21167)"

5'UTR /note="ecotype: Columbia"

CDS 1..3205

206..1216 /note="unknown protein"

/codon_start=1

/product="Atg1975080/P9E10_7"

/protein_id="AA157684.1"

/db_xref="GI:18086461"

/translation="MTSDGATSTSAAAAAAAAAARRKPSMERENRRRRRAVA AKITVGLRAQGDYVLPKQINNEVLKALCEAGWVVEEDGTTKCKPLPGEIAGTS SRVTPSQNQSPILSSAFQSPISYQVSPSSSPSPSGEPNNNSSTFPFLRNG IPSGLPSLRISNSCVTPPLISPTSPKLPKLPNWSIAKOSMAIAKOSMAFVYFPG VSRAPSPTRHQPHPTPATIPEDDESQSVDSGHWISFQKFAQOQFEASVMTPTPF NIIVKPAQOQSPNTAIPQETGQSSSEFPENSOVKPEMERIHVDGHEDELITGNKKA RG"

BASE COUNT 3'UTR 390 a 344 c 306 g 436 t

ORIGIN

Alignment Scores:

Pred. No.: 7.55e-73 Length: 1476

Score: 1755.00 Matches: 334

Percent Similarity: 99.70% Conservative: 1

Best Local Similarity: 99.40% Mismatches: 1

Query Match: 99.43% Indels: 0

DB: 8 Gaps: 0

US-09-995-938A-7 (1-336) x AY065049 (1-1476)

QY 1 MetThSerAspGlyValThrSerThrSerAlaAlaAlaAlaAlaAlaAla 20

DB 206 ATGACTCTCGATGAGACTACGTGACATCAGCAGCTGAGCTCGGCGGCGAGCG 265

QY 21 AlaArgArgLysProSerTPARArgLysArgLysAsnAsnArgArgArgArg 40

DB 266 GCGAGAGAGAGCGCGTCGTGAGAGAGAGAGAGATATCGAGAGAGAGAGAG 325

QY 41 ArgAlaValAlaAlaLysIleTyrThrGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro 60

DB 326 AGAGCTGTGAGCGGAGAGATATCAGCTGAGCTTACAGCTTAAATTTGGCT 385

QY 61 LysHisCysAspAsnAsnGlnValLeuLysAlaLeuCysValGlnAlaGlyTyrPvalVal 80

DB 386 AAACATTTGATATATATATAGATCTTAAAGCTTTGTGTGAAGCTGTGGTGTGT 445

QY 81 GlnGluAspGlyThrThrTyrArgLysGlyCysLysProLeuProGlyGluIleAlaGly 100

DB 446 GAAAGAGATGTACTACTATTCGCAAGAGATGCAAGCTTACTGTGATAGCTGG 505

QY 101 ThrSerSerArgValThrProTyrSerSerGlnAsnGlnSerProLeuSerSerAlaPhe 120

DB 506 ACTTCATCTCGAGTATCATCTCATTCATCACAAGACAGAGCCCTTTCATCAGCCTTT 565

QY 121 GlnSerProLysProSerTyrGlnValSerProSerSerSerSerPheProSerProSer 140

DB 566 CAAAGTCCATCCCATCTTACCAAGATGAGCCGCTTCTTCATCATTCGCCAGTCTCT 625

QY 141 ArgGlyGlyProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAsnGlyGly 160

DB 626 CCGGAGAACCAATATCAACATGTCTCTCAATTTCTCCCTTCCCAAAATGTGGC 685

QY 161 IleProSerSerLeuProSerLeuArgLysSerAsnSerCysProValThrProVal 180

DB 686 ATTCCTTTCTTCTCTTCCCTCCAGATATCAAGAGTGTGACAGTATCCCAAGTCTC 745

QY 181 SerSerProThrSerTyrAsnProLysProLeuProAsnTrpGlnSerIleAlaLysGln 200

DB 746 TCATTCGCGAGCTTCAAGAACCGAAGCCGTGCTTACCTGAGATATATGCTAACCAA 805

QY 201 SerMetAlaIleAlaLysGlnSerMetAlaSerIleAsnTyrProPheTyrAlaValSer 220

DB 806 TCCATGGCACTTCTCAACATCAATGAGCTCTTTTAATATATCTTCTATGCGGTTCT 865

Oy	22	ALAProAlasSerProThrHisArgIleGlnPheHisSerTrpLeuAlaThrLilePerGluLucys	240
Db	866	GCACTGGTGTAGTCACACACATCGCACAGTTTACTATTACCCCGGCTACTATACCTGAATGT	925
Oy	241	AAPgluSeAapSseSerSerThrValAspSerGlyHisTrpLileserPheGlnLylsPheAla	260
Db	926	GATGAGCTGTGACTCTTCCACTGTTGATTTCTGGTACTTGGAATAGCTTTCAAGAAGTTTGCA	985
Oy	261	GInGInGInLPropheSerAlaSerMetValProThrSerProThrPheAsnLyuVallys	280
Db	986	CACAACAACGCATCTCTCGCCTCATGTGTCGCCAACCTCTCTTAACCTTCAATCTGTGTAA	1045
Oy	281	ProHlaPProGInGInMerSerProSerThrAlaalaPheGInGInLIGLyGInSerSer	300
Db	1046	CCTGGCCCTCAGCAGATGTCTTCCAATACTGCTGCTTCCACAAGATTTGTCAAAAGCTCT	1105
Oy	301	GIUpheLysPheGluAsnSerGlnValIysProTrpGluGlyGluArgGIlleHisApVal	320
Db	1106	GAGTTTAAATTGGAGAAATAGCCCAATTAAACCTCGGAAGAGAGAGATACATGATGTG	1155
Oy	321	GlyMetGluAspLeuGlnLeuThrLeuGlyAsnGlyIysValaArgGly	336
Db	1166	GGTATGGAGAGATCTTGAGCTTACACTTGGAATGGGAAGCGCTCGTGT	1213
RESULT 4			
LOCUS	AF494338	2687 bp	DNA linear PLN 20-MAY-2002D
DEFINITION	Arabidopsis thaliana BZR1 protein (BZR1) gene, complete cds.		
ACCESSION	AF494338		
VERSION	AF494338.1	GI:20270970	
KEYWORDS	Arabidopsis thaliana (chale crese)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;		
TITLE	roside; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
JOURNAL	1 (bases 1 to 2687)		
MEDLINE	Wang,Z.-Y., Nakano,T., Gendron,J., He,J., Chen,M., Vafeados,D.,		
PUBMED	Yang,Y., Fujikawa,S., Yoshida,S., Asami,T. and Chory,J.		
REFERENCES	Nuclear-localized BZR1 mediates brassinosteroid-induced growth and		
AUTHORS	feedback suppression of brassinosteroid biosynthesis		
TITLE	Dev. Cell 12 (4), 505-513 (2002)		
JOURNAL	Submitted (22-MAR-2002) Plant Biology, The Salk Institute for		
BIOLOGICAL STUDIES	Biological Studies, 10010 N. Torrey Pines Rd., La Jolla, CA 92037,		
USA			
FEATURES			
source	Location/Qualifiers		
	1..2687		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:3702"		
	/chromosome="1"		
	/map_between ngaili and ARArPasase"		
	/note="ecotype: Columbia"		
	1..2687		
	/gene="BZR1"		
	1..880		
	/gene="BZR1"		
	join(881..1369,1714..2687)		
	/gene="BZR1"		
	/product="BZR1 protein"		
	join(1103..1369,1714..2457)		
	/gene="BZR1"		
	/codon_start=1		
	/product="BZR1 protein"		
	/protein_id="AAAI8490.1"		
	/db_xref="GI:20270971"		
	/translation="MTSGATSTSAAAAAAAAAARRKPSRRENNRRERRRRRAVA		
	AKIYSLAQGDVYNLRKHCDNNEVTALKALCVENGWVEBDGTYRGCKPLGLAGTS		

[illegible]


```
mRNA
CDS
    join(<4711..4845,5272..>6345)
    /gene="F9E10.3"
    /join(4711..4845,5272..6345)
    /gene="F9E10.3"
    /product="unknown protein; 4711-6345"
    /protein_id="AAG51930.1"
    /db_xref="GI:12333904"
    /translation="MAVRKEKVPFRCGIAIAIVGIFGCVCTIIIPNDFVNFSS
KVASASCESEPERVAFAPAIISEKNGELKQVSDTEKRLAEQKEVIKAGPGTV
GLQINPTVAIDESANPRLAKLEKVAIVKEIIVLVANNVPMLEVOIASVKVGVG
NYLVPLDLSLESCGNEVAYYKRDNDIAIDVYKRSRSDVGLKFRVLRESLQIG
YGLVLSVDIVLQNPGRHLYRSDVSMGSDGHNNMTAGSNDVFPDPTMRSTVTV
NRIWVRNSGFFYLRLPTLPSIELDRVTDTLSKSGMDQAFVNOILTPSPHGTGLYA
SKVMVDYEFEMNSKVLFTVKQDEMKKLKPVIIHNNYHSDKLERMQAAVEFYNGQ
DALDRFRDGS"
    7482..9019
    /gene="F9E10.4"
    /join(<7482..7655,7907..>9019)
    /gene="F9E10.4"
    /join(7482..7655,7907..9019)
    /gene="F9E10.4"
    /codon_start=1
    /product="unknown protein; 7482-9019"
    /protein_id="AAG51917.1"
    /db_xref="GI:12333891"
    /translation="MAGRRIQQLRGSRIAIAIFVILIGCVSVLPNGFNGSGS
LNIERISKSTSDGLASCESESRVMLSDSIIISVNAELKQVRELTEKRLAE
QETENAKQVLVIGSEIKAGPFGVNTLPVPSVNPRLAKLEKVAIVKEI
VVLANSNVKPMLEIOIASVKRQVIONLIALDSDMSFCSEKVPVYKRDPAVIM
VNGSGNHAVSGIKFRYLREPLQVGLVLSVDIVLQNPGRHLREPLHSDVESDQ
DNNTAGFNVPPERPSSGMARYATNRIWVRNSGFFYLRLPTLPSIELDRVADLDSK
EANDQAFNQQLTPSPHGTGLYASRVRVDMVEFMSKVLFTVRNQLKLPVI
VHANNYHSDKLERMQAAVEFYNGQDALDSFPDGS"
    9598..12387
    /gene="F9E10.5"
    /join(<9598..10007,10087..10712,10801..11195,11275..11389,
11467..11516,11613..11678,11776..11853,11944..11210,
12221..12387)
    /gene="F9E10.5"
    /join(9598..10007,10087..10712,10801..11195,11275..11389,
11467..11516,11613..11678,11776..11853,11944..12120,
12221..12259)
    /gene="F9E10.5"
    /codon_start=1
    /product="unknown protein; 9598-12259"
    /protein_id="AAG51921.1"
    /db_xref="GI:12333895"
    /translation="MOTLPSESEVTLGSSNAPPVLRSGDDVDIDFGDVFGPPKRR
SKYTSNEVTRHSFSESALRRDYIVDGDLLPQDEKVFGEIDSSVRRRTTDDFPDD
IFRVNSSSSLPGRSILSPAKPBESSGTSPPQSLPAKATEIPTFLAATRLANKK
ETVSSPILSTRSKADVSTAKSDDCDDPOVFTVGKROFHSIYKKNQGVPLV
IWSRSLSSMSKABETTPVPLSDYRKTSVYKKEGSGKLSGLKXVKTSTKR
PGVOTKEKETETDLKSGQAFVYSKABEAVKPLDSVESQARSQSKAEATTVKRL
HSIFHEDEKQDEKIVSEREVKRGSKAKATRSFTESRTKKSQGTSLDSSPLPD
KSSFPASSASPEVKGQGVKSDPVIKFKGSGVAGSGSLQSSSMRAREPKDI
IHDSNAKEVNIIPDOOKSSTPDI PANNRKQPSQSTOKSDRESNRYAKPEKDI
ERQESTTTTSEDIDEFPVNFVEDITDDENKMEANQDAEIKIINDIKIRMSG
KSGNIRSLSTLQYILMSGGKVPPLMDIEGNAVRKSGVORALLIHPKLOOKGAS
AANKYMAEKVFEILOEMDHPNLTGPV"
    complement(14275..14526)
    /note="DIT1 Non-autonomous DNA transposon [a
consensus]"
    14940..15720
    /gene="F9E10.6"
    /note="similar to putative DNA-3-methyladenine glycosylase
I (TAG I) GB-P05100 [Bacteriella coli]"
    /join(<14940..15100,15179..15320,15400..>15720)
    /gene="F9E10.6"
    /join(14940..15100,15179..15320,15400..15720)
    /gene="F9E10.6"
    /codon_start=1
    /product="putative DNA-3-methyladenine glycosylase I;"
```

```
14940-15720"
    /protein_id="AAG51925.1"
    /db_xref="GI:12333899"
    /translation="(MVIDQTP)IVLVLPHDEMGVPRVDKDLFELVFSQALAEPSWP
SILRRDPRKLFEPDPSAIAGTEKRLMSLRVNGCLILSEOKLRAIVENKSKVLKY
KQEGFSNRYCMWPFVNHKPLRANGRVGRVPRVPSKPAEYISKDMNORGFRCVPTVMY
SFIQNSGVVNDHITACRYORQVETRETKSHETETKLDHSLPV"
    complement(16304..17716)
    /gene="F9E10.7"
    /join(<16304..17038,17392..>17716)
    /gene="F9E10.7"
    /complement(join(16304..17038,17392..17658))
    /gene="F9E10.7"
    /codon_start=1
    /product="unknown protein; 17658-16304"
    /protein_id="AAG51929.1"
    /db_xref="GI:12333903"
    /translation="(MISDGA)STGSAIAAAAAAAAAARRKPSRENNRRRRARRAVA
AKIYTGIRAGQDYNLPRKCDNNEVLKALCYEAGVNVEDGTTTRKPLFELAGTSRV
TPYSQVQSPSLSSAFOSPIPSYOVSPSSSPSSRDEPNMNSSTFPFLRNGI PS
SLPSLRISNCPVTPVSPSTSKNPKPLPMWESIAKOSMAIAKOSMAFNYPVAVSA
PASPTRHQFTPATIIECDSDSSTVLSGHWISFQKFAQQQPSASVPTSPFNLV
KPAQQMSPTTAFOEISGSEKFEKSVQKPMWEGRIHDVGMEDLELTNGKXRG"
    complement(18764..18835)
    gene
```

```
Alignment Scores:
Pred. No.: 1,276-66 Length: 97263
Score: 1664.00 Matches: 334
Percent Similarity: 74.06% Conservative: 0
Best Local Similarity: 74.06% Mismatches: 2
Query Match: 94.28% Indels: 116
DB: 8 Gaps: 1

US-09-995-938a-7 (1-336) X AC013258 (1-97263)

QY 1 MetThSernePGLYAaThrSerThSerRLAaLaLaLaLaLaLaLaLaLa 20
Db 17658 ATGACTTCGATGATGAGCTAGCTGACATCAGACCTCAGCTGGCGGCGACGCG 17599

QY 21 ALaArggLySPoseRTPaRGyLaRGyLuBenaMaYgaYgaYgaYgaYgaYg 40
Db 17598 GCGAGAGAGAACCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 17539

QY 41 ArgAlaValAlaLaYleIeYThRgLYeuaRgLaInGIyaSPYyAsnLeuPro 60
Db 17538 AGAGCTTACCTCGAAGATATACACTGGGCTTAGACTCAAGGTGATTATTTGCT 17479

QY 61 LyehiScyabPbenaBngLuvallLeuYAlaLeuCYsValGIuAlaGIYTryValaI 80
Db 17478 AACATTGTGATATATATGAAAGCTCTTGAAGCTCTTGTGTAAGCTGTTGGTCTT 17419

QY 81 GluGluaspGLThrThrYfayRgLy- 89
Db 17418 GAAGAGATGATCTACTATTCGAA-GGTGAAGACTTCTCCATTTTTCAGATCTGA 17360

QY 89 - 89
Db 17359 GCTTGTATTATGATGTTTGTGATGTTGAATCGAATTCGATTTCATTTGCTTA 17300

QY 89 - 89
Db 17299 AATGGTTGAATCTGAGAAATTGAGGATTTTCTCAAGTGAATTGAAATCATCAGAA 17240

QY 89 - 89
Db 17239 TATGATGATGATGATTTCTCAAGTGAATTATGAGTTTCTTTCTTAATTTAGAGTTA 17180

QY 89 - 89
Db 17179 TTATGTGATGCTAAAGCTTAATCTTTATGATGATGATGATGATGATGATGATGAT 17120

QY 89 - 89
```


Db	17119	IGTGTTCCTTTGCTTACCTGTGATGTATGATGTTGGATTGCTTATGTTTGGCTTTTG	17060
Oy	90	-----GlyCybelysProleupProgllyGlu1lealaglyThrsSerArval	105
Db	17059	TTGGAGTATCGAGGATGCAACCTTTTACCTGGTAGATACCTGGAGCTTATCTCCAGTA	17000
Oy	106	ThrProTysSerSerGlnAsnGlnSerProLenseSerAlaPheGlnSerProilePro	125
Db	16999	ACTCCATTTATCATCGAACCAAGAGCCCTTTTCATCACCTTTCAAACTCCATCCCA	16940
Oy	126	SerTyrGlnValSerProSerSerSerSerPheProSerProSerArvalGlyUpProan	145
Db	16939	TCTTACCAAGATTAGCCCTCTTCTTATCATTCGCCAGTCTCTTCGGCGGTGAACCAAT	16880
Oy	146	AsnAsnMetSerSerThrPhePheProPheLeuArgAsnlyGly1leProSerSerleu	165
Db	16879	AACAACATGCTCTTACATCTTCTCCCTTTCCTCAGAAATGATGGACATTCCTTCTCTT	16820
Oy	166	ProSerleuAglileSerAsnSerCyProValThrProProValSerSerProThrSer	185
Db	16819	CTTTCCTCCAAATCTCAAAACAGTTGTCAGATTACCCCAACGGATCTCATATCGACATTCT	16760
Oy	186	LyAsnProLybProLeuProAntTrrGlySer1leAlaLybGlnSerMetAla1leAla	205
Db	16759	AAGAACCAGAACCTGTCCTTACCTGGAAATCTATCCCTTAAGCAATCATGSCCATTTGCT	16700
Oy	206	LyGlnSerMetAlaSerPheAsnTyrProPheTyrAlaValSerAlaProAlaSerPro	225
Db	16699	AAACATCAATAGGCGCTCTTTTAATTATCTTTTCTATGCGGTTTCTGACCTGTAATCGG	16640
Oy	226	ThrThiAlaTgH1egInPheH1stH1leuAlaThr1leProGlyCybArgGlySerArpSer	245
Db	16639	ACACATGCGCCACAGTTTCAATCCCGGCTACATACCTCAATGATGATGATGCTGACTCT	16580
Oy	246	SerThrValAspSerArgLyH1atTrr1leSerPheGlnLybPheAlaGlnGlnGlnPhePhe	265
Db	16579	TCCACTGTGATTTCTGTGATTCGATAGATAGCTTTTCAGAACTTTGCACACACAGCCATTC	16520
Oy	266	SerAlaSerMetValProThrSerProThrPheAsnLeuValLybProAlaProGlnGln	285
Db	16519	TCTGCGCTTATAGTGCCAACTCTCTACCTCAATCTTGTGAACCTGGCGCTCAGCAG	16460
Oy	286	MetSerProAntThrAlaAlaPheGlnGlu1leGlyGlnSerSerGluPheLybPheGlu	305
Db	16459	ATGCTCTCAATACTGCTGCTCTTCCAAAGATATGTCCTAAAGCTCTGATTAATTTGAG	16400
Oy	306	AsnSerGlnValLybProTrrGlyGlyGluArg1leH1stAspAlaGlyMetGluAspLeu	325
Db	16399	AATAGCCCAAGTTAAACCTCGGAGAGAGAGATGATCATGATGTGGATAGAGATCTT	16340
Oy	326	GluLeuThrLeuGlyAsnGlyLybAlaArgGly1y 336	
Db	16339	GAGCTTACCTTGGAAATGGGAAGCTCTGCGT 16307	
RESULT 6			
AY065041			
LOCUS	AY065041	1340 bp	mRNA linear, PLN 07-JAN-2002
DEFINITION	Arabidopsis thaliana Atlg19350/F18014_4	mRNA, complete cde.	
ACCESSION	AY065041		
VERSION	AY065041.1	GI:18086445	
KEYWORDS	FLI cDNA.		
ORGANISM	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Chauk, R., Chen, H., Kim, C. J., Koesema, E., Meyers, M. C., Banb, J., Bower, L., Cernulnik, P., Chang, E., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Miranda, M., Natsuka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Tortum, M., Wu, H. C.,		

Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1340)
Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (30-NOV-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN Arabidopsis Full-Length cDNA). : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P. X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Phan, P. K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.
Chen, R. (SSP/salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as pIs.
Location/Qualifiers
1. 1340
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF105-03-B11(R21175)"
/note="ecotype: Columbia"
1. 137
138..1145
/note="unknown protein"
/codon_start=1
/product="At1g19350/F18014_4"
/protein_id="AA57677.1"
/db_xref="GI:18086446"
/translation="MTSDATSTYLSAAAAAAMATRRKPSRENNRRRRRAVA
KITYGARAQGNVYLPRKCDNNELYLKASEAGVVEDETTTKRGKRLPGDAGSS
RAIPFSHNOSEPLSTFDPSFLISTQVSSSSPPSRVGDPPNISTIFELKNGIP
SLSPLPRLINSAPVTPPVSSPTSRNRPPLPTWSSFTKQSWMAKQSSFLNPFYAV
SAAPSSTHROFAPATIPEDCSGDSITVDSGMISFOKXAQCOOPASASVPSPTFN
LVNPAEQQLSPNTAAIQEIGSSSEFFKENSQVPEWGERIHDVAMEDLELTNGKAKH
S"

```

Db      138 ATGACGCTGACGAGCACTCCACGTC---GCTGAGCTGACGACGACCATGCGC 194
Qy      21 AlAargArglyseProSerTPArgGluArgGluAuaAnaAgaArgArgGluArgArg 40
Db      195 ACAGAGGAGAAACCGTCGTGAGAGAGAGAGAAACATCGAGAAAGAGACGGCGGAGA 254
Qy      41 ArgAlaValAlaAlaValleTyrThrGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro 60
Db      255 AGAGCTGTTCGGGAGAGATTATATACGTCTTGAAGCTCAAGGTAATCAATCTTCCA 314
Qy      61 LyHHisCyAaPnaAnaGluValLeuLysValLeuCyAaValGluAlaGlyTyrValVal 80
Db      315 AAACATTGTGACAAACATAGAGTCTTAAGGCTCTTTGTTGTAAGCTGAGTGGGTTGT 374
Qy      81 GluGluAaPglYThrThrTyrArgGlyGlyCyAaLysProLeuProGluGluAlaGly 100
Db      375 GAGAGAGACGGAATCTTATCCGACAGGACAAAGCCTTAAGTGGTGAATGGCTGGA 434
Qy      101 ThrSerSerArgValThrProTyrSerSerGlnAaGlnSerProLeuSerSerAlaPhe 120
Db      435 TCATCTTCTCGAGCACTCTTACTCTTCCATACCAAGTCCCTTTCTTCCACTTTT 494
Qy      121 GlnSerProLleProSerTyrGlnValSerProSerSerSerSerPheProSerProSer 140
Db      495 GATAGCCCCATCTTATCTTACCAAGTCAAGTCTTCTTCTTCAATCCGAGTCTTCT 554
Qy      141 Arg---GlyGluProAaAnaAnaMetSerSerThrPhePheProPheLeuArgAsnGly 159
Db      555 CGAGTGGTATCCACCAAT-----ATCTCCCAATCTTCCCTTCTTCCAGGAATGGT 608
Qy      160 GlyYlleProSerSerLeuProSerLeuArgLleSerAaSerCyAaProValThrProPro 179
Db      609 GGAATTCCTTCATCGCTTCTCCCACTAGATCTCAACAGTGCCTCTGTCATCCACCA 668
Qy      180 ValSerSerProThrSerLysAaAnaProLysProLeuProAaThrPglLysLleAlaLys 199
Db      669 GTGTCACTCCCACTCTAGAACCCCAACCAATGCTGCTTGGAAATCTTTTACCAAA 728
Qy      200 GlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPheAnaTyrProPheTyrAla 218
Db      729 CAATCCATGTCANCGGCTGTAACATGATGATCTTCTTGAACATACCGCTTTATGCG 788
Qy      219 ValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrLlePro 238
Db      789 GTGTCTCCACTGCGCACTCTCACTCATCATCGCAGATCTCCGCGCTACTATACCT 848
Qy      239 GlyCyAaPglLysAaPnaAnaSerSerThrValAspSerGlyHisTyrPllSerPheGlnLys 258
Db      849 GAATGTATAGTCTGACTCTTCCACTGTGATTTCTGTCTTGGATTAAGCTTTCAAAAG 908
Qy      259 PheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsnLeu 278
Db      909 TTGGCAACAACAAGCCATTTCTGCTCTTATGATGCAACCTGCTTCACTTCAATCTC 968
Qy      279 ValLysProAlaPnaAnaSerProAnaThrAlaAlaPheGlnGluLleGlyGln 298
Db      969 GTGAACCTGACACACAGCAATTTGTCTCAAAACAGAGCAATCCAAGATTTGGTCAA 1028
Qy      299 SerSerGluLysPhePheGluAaAnaSerGlnValLysProTyrPglLysGluArgLleHis 318
Db      1029 AGCTCCAGATTTAAGTTTGAACACAGCAAGTTTAAGCATGGGAAGGAGAGATCCAT 1088
Qy      319 AspValGlyMetGluAaPnaAnaGluLeuThrLeuGlyLysGlyLysVala 334
Db      1089 GATGTGCTATGAGAGATCTAGACTCAAGCTTTGAAATGTAAAGCT 1136

```

```

RESULT 7
AF134217 1257 bp mRNA linear PLN 01-JAN-2000
LOCUS Arabidopsis thaliana 107 protein (107) mRNA, complete cds.
DEFINITION
AF134217
ACCESSION
AF134217.1 GI:6651068
VERSION

```

```

KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (chale crees)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1257)
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
2 (bases 1 to 1257)
AUTHORS
Kwon,H.-B.
TITLE
Molecular Cloning and Characterization of a Gibberellin-responsive
Gene from Arabidopsis thaliana
JOURNAL
Unpublished
AUTHORS
Kwon,H.-B.
TITLE
Direct Submission
JOURNAL
Submitted (10-MAR-1999) Bioresources, National Institute of
Agricultural Science and Technology, Seodun-Dong 250, Suwon,
Kyunggi-do 441-707, Korea
FEATURES
source
1..1257
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
1..1257
/gene="107"
/feature="responsive to gibberellin in Arabidopsis mutant ga3"
55..1062
/codon_start=1
/product="107 protein"
/protein_id="AF22161.1"
/db_xref="GI:6651069"
/translation="MTSDGATSTSAAMAAAMATRRKPSWRENNRRRRRAVA
KIYTGUAQCNVLPKHCNNBYLKALISEGAWVEEDGTYRKGKPLPCDMAGSS
RATPYSDNOSPSTFSDSPILSYQVHSSSPSPSPVCPDHPNISTFPPLRNGIP
SSLPRLISNAVTPVPSPTSPNPKPLPWPESFTQSMMAAKQSWTSLNPRYAV
SAPASPTHHQFAPATIPEDCSDDSTYDSGHRISFQKFAQQQPFASMPSTSTFN
LVKPAFQPLFPNTAIGTQSSFEKFNQVKEBERLHDVAMEDELTLGNGKAH
S"
BASE COUNT 325 a 290 c 273 g 369 t
ORIGIN
Alignment Scores:
Pred. No.: 8,83e-62 Length: 1257
Score: 1513.50 Matches: 296
Percent Similarity: 91.07% Conservative: 10
Best local Similarity: 88.10% Mismatches: 25
Query Match: 85.75% Indels: 5
DB: Gaps: 4
US-09-995-938A-7 (1-336) x AF134217 (1-1257)
Qy      1 MetThrSerArgGlyAlaThrSerThrSerAlaAlaAlaAlaAlaAlaAla 20
Db      55 ATGACGCTGACGAGCACTCCACGTC---GCTGAGCTGACGACGACCATGCGC 111
Qy      21 AlAargArglyseProSerTPArgGluArgGluAuaAnaAgaArgArgGluArgArg 40
Db      112 ACAGAGGAGAAACCGTCGTGAGAGAGAGAGAAACATCGAGAAAGAGACGGCGGAGA 171
Qy      41 ArgAlaValAlaAlaValleTyrThrGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro 60
Db      172 AGAGCTGTTCGGGAGAGATTATATACGTCTTGAAGCTCAAGGTAATCAATCTTCCA 231
Qy      61 LyHHisCyAaPnaAnaGluValLeuLysValLeuCyAaValGluAlaGlyTyrValVal 80
Db      232 AAACATTGTGACAAACATAGAGTCTTAAGGCTCTTTGTTGTAAGCTGAGTGGGTTGT 291
Qy      81 GluGluAaPglYThrThrTyrArgGlyGlyCyAaLysProLeuProGluGluAlaGly 100
Db      292 GAGAGAGACGGAATCTTATCCGACAGGACAAAGCCTTAAGTGGTGAATGGCTGGA 351
Qy      101 ThrSerSerArgValThrProTyrSerSerGlnAaGlnSerProLeuSerSerAlaPhe 120

```

```

Db      352 TCATCTTCGAGCACTCTTACTCTTCCATACCAAGTCTCTTTCACACTTT 411
Qy      121 GlnSerProIleProSerTygInValSerProSerSerSerPheProSerProser 140
Db      412 GATAGCCCATCTTATCTTACCAAGTCAAGTCTCTTCTTCTTCAATCCGAGTCTTCT 471
Qy      141 Arg---GlyIupProMetMetSerSerSerTherPhePheProPheLeuArgAngly 159
Db      472 CGAGTTGGTGCATCACAACAT-----ATTCACACATCTTCCCTTCTCAGAGATGGT 525
Qy      160 GlyIleProSerSerLeuProSerLeuArgIleSerAmsSerCysProValThrProPro 179
Db      526 GGAATTCCTTCATCGCTTCTCCATGATGATCAAAAGTCAAGTCTCTCTGTCATCCACCA 585
Qy      180 ValSerSerProTherTybAnProlYsProlLeuProAnThrPgluSerIleAlaYs 199
Db      586 GTGTCACTCCCACTTCTAGAAACCCCAACCAATTCCTACTGGAAATCTTTTACCAAA 645
Qy      200 GlnSerMetAlaIle---AlaYsGlnSerMetAlaSerPheAnTyrrProPheTyra 218
Db      646 CAATCATATTCATGCTGCTGTAACAGTCAATGACTTCTTTGAACCTAACCGTTTATGGC 705
Qy      219 ValSerAlaProAlaSerProTherHsArgHsGlnPheHsThrLeuAlaThrIlePro 238
Db      706 GTGTGCGACCTGCCAGTCTCTACTCATTCAGCCAGTTCCTCGGCTCATATACCT 765
Qy      239 GluCyAspGlnSerSerSerSerTherValAspSerGlyHsIleTrpIleSerPheGlnYs 258
Db      766 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 825
Qy      259 PheAlaGlnGlnGlnPhePheSerAlaSerMetValProTherProTherPheAnleu 278
Db      826 TTTCACAAACAACAGCATTCTCTGCTCTATAGTGCACCACTCGGCTTCACTCAATCTC 885
Qy      279 ValIysProAlaProGlnGlnMetSerProAnThrAlaAlaPheGlnIleGlyGln 298
Db      886 GTGAACCTGCGACACAGCAATGTTTCCAAACACAGCAATCCAAAGATGGATTCGA 945
Qy      299 SerSerGlnPheIlePheGlnLeuSerGlnValIysProTrpGlnIleGlyArgIleHs 318
Db      946 AACTCCGAGTATTAAAGTTTGAAGAACCAAGTAAAGCAATGGAAGGAGAGATTCAT 1005
Qy      319 AppValGlyMetGlnAspLeuGlnLeuThrLeuGlyAspGlyValAla 334
Db      1006 GATGTGGCTATGAGAGATTAAAGCTCAGCGTTGGAAGATGTAAGCT 1053

```

RESULT 8
AY086340
LOCUS AY086340 1359 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 24157 mRNA, complete sequence.
ACCESSION AY086340
VERSION AY086340.1 GI:21405050
FLI CDNA
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote, **Vitridiplantae**; **Streptophyta**; **Embryophyta**; **Tracheophyta**; **Spermatophyta**; **Magnoliophyta**; **eudicotyledons**; **core eudicots**; **rosoids**; **eurosidids II**; **Brassicales**; **Brassicaceae**; **Arabidopsis**.
REFERENCE 1 (bases 1 to 1359)
AUTHORS Haas, B.J., Volkovskiy, N., Town, C.D., Troupkan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE 22088475
PUBMED 12093376
REFERENCE 2 (bases 1 to 1359)
AUTHORS Brover, V., Troupkan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1359)

AUTHORS Brover, V., Troupkan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated, less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES
source
 1..1359
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="24157"
 112..1119
 /codon_start=1
 /product="unknown"
 /protein_id="AA064408.1"
 /db_xref="GI:21592457"
 /translation="MTSDGATSTSAATAAATRRKPSRERENRRRRRAVAA
 KIVTGLRAQGNLPEKCDNNEVLALCEAGMVEEDGTYRKGRPLPGDAAGSS
 RATPSSHNOSPLSTFDSPISTYOVSPSSSPSPSRGDDPNISTIPFLNGGIP
 SSLPLRLISNSAPVTPVPSPTPSRSKSLPTMESITKQMSMAAKSMTSLNPFAY
 SNAPPTTHRQFHATTPBCDSBSSTYDSCHWISFQFAQOOPFSASNVPTSPFN
 LVKPAPOUSPTPAATIQETQSSSEKPFENSVQKFWGERIHDAVMEDELITGNRAH
 S"

CDS
 356 a 317 c 285 g 401 c
 BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 1,19e-61 Length: 1359
 Score: 1511.50 Matches: 296
 Percent Similarity: 91.07% Conservative: 10
 Best Local Similarity: 88.10% Mismatches: 25
 Query Match: 85.64% Indels: 5
 DB: 8 Gaps: 4

US-09-995-938a-7 (1-336) * AY086340 (1-1359)

```

Qy      1 MetTrSerAspGlyAlaTherSerTherSerAlaAlaAlaAlaAlaAlaAla 20
Db      112 ATGACGCTGACGAGCAAGTCAAGTCA---GCTGACGCTGACAGCAGCGATGGCG 168
Qy      21 AlaArgArgIysProSerTrpArgGlnArgGlnAsnAsnArgArgArgArgArg 40
Db      169 ACGAGAGAGAAACCGTCGTGAGAGAGAGAGAGAGAACAAATCGGAGAGAGCGCGGAGA 228
Qy      41 ArgAlaValAlaAlaIlysIleTrpGlyLeuArgAlaGlnGlyAspTyraAnleuPro 60
Db      229 AGAGCTGTTGCGCGCAAGATTATCTGCTTAAAGCTCAAGGTAACATACTTCCCA 288
Qy      61 LysHsIcyAspAsnAsnGlnValLeuIysAlaIcyValGlnIleArgIleTrpValAla 80
Db      289 AACATTTGTGACACCAATGAGGTTCTTAAGGCTCTTGTTCAGACCTGGTGGTGGT 348
Qy      81 GluGluAspGlyThrTherTyArgIysGlyCysIysProLeuProGlyGluIleAlaGly 100
Db      349 GAAAGAGACGAACTACTTATGCAAGGAGACACAGCCCTTACCTGCTGATCAGCTGGA 408
Qy      101 ThrSerSerArgValThrProTySerSerGlnAenGlnSerProLeuSerSerAlaPhe 120

```

```

Db      409  TCATCTTCTCGACGACCTCTTACTCTTCCATAACCAAGCTCTTCTTCCACTTT 468
Qy      121  GlnSerProIleProSerTyGlnValSerProSerSerSerSerPheProSer 140
Db      469  GATAGCCCATCTTATCTTACCAAGTCAGTCTTCTCTTCTTCACTCCGAGTCTTCT 528
Qy      141  Arg---GlyGluProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAngly 159
Db      529  CGAATTGGTATCCACACAAAT-----ATCTCCAAATCTTCTCTTCTTCCAGGAAATG 582
Qy      160  GlyIleProSerSerLeuProSerIleuArgIleSerAsnSerCysProValThrPro 179
Db      583  GGTATTCCTTATGCTGCTTCTTCCACTTGAATTCMAACAGTCTCTCTTCACTCCAC 642
Qy      180  ValSerSerProThrSerLeuAsnProLysProLeuProAsnTrpGluSerIleAla 199
Db      643  GTGTATCTCCCACTTCTAGAACTCCAAACCATTTGCTTGGGAACTTTTACCAA 702
Qy      200  GlnSerMetAlaIle---AlaIleGlnSerMetAlaSerPheAsnTrpProPheTy 218
Db      703  CAATTCATGTCACAGGCTGCTMAACAGTCAATGACTTTTGAACCTACCCGTTTAT 762
Qy      219  ValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThr 238
Db      763  GTGTCTGCACTGCTGCACTCTTACTATCATGCGCAATTCATGCTCGGCTACTAT 822
Qy      239  GluCysAspGluSerAspSerSerThrValAspSerGlyHisTrpIleSerPheGln 258
Db      823  GAATGTATGATGCTGACTCTTCCACTGTTGATTTGCTGATGATGATGATGATG 882
Qy      259  PheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPhe 278
Db      883  TTGTCACAAACAAACGCAATCTCTGCGCTTGTGGCCAACTGCGCTTCACTCA 942
Qy      279  ValIleProAlaProGlnGlnMetSerProAsnThrAlaAlaPheGlnGlnGln 298
Db      943  GTGAAACCTGCACACGCAATGCTTCCAAACAGACGAAATCCAAAGATGATG 1002
Qy      299  SerSerGluPheLysPheGluAsnSerGlnValLysProTrpGluGlyLysArg 318
Db      1003  AGCTCCGAGTTAAGTTTGAACAACGCAATTAAGCATGGAAGGGAAGAGATCAT 1062
Qy      319  AspValIleMetGluAspLeuGluLeuThrLeuGlyAnglyLysAla 334
Db      1063  GATGTGCTATGAGAGATCTAGAGCTCAGCTGGAATGTAAGCT 1110

RESULT 9
AC025808 120977 bp DNA linear PLN 11-OCT-2000
LOCUS Genomic sequence for Arabidopsis thaliana BAC F18014 from
DEFINITION chromosome 1, complete sequence.
AC025808
VERSION AC025808.8 GI:7636235
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 120977)
Shim, P., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C.,
Khan, S., Kim, C., Altfeld, H., Bel, O., Chen, C., Chou, J., Choi, E.,
Conn, L., Conway, A., Gonzalez, A., Hansen, N., Howing, B., Koo, T.,
Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S.,
Mukharbek, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J.,
Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G.,
Federici, N., Theologis, A. and Ecker, J. R.
Genomic sequence for Arabidopsis thaliana BAC F18014 from
chromosome 1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 120977)
AUTHORS Ecker, J. R.

```

```

TITLE Direct Submission
JOURNAL Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center,
AUTHORS Department of Biology, University of Pennsylvania, 38th Street and
REFERENCE Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
3 (bases 1 to 120977)
Ecker, J. R.
TITLE Direct Submission
JOURNAL Submitted (22-APR-2000) Arabidopsis thaliana Genome Center,
AUTHORS Department of Biology, University of Pennsylvania, 38th Street and
REFERENCE Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
4 (bases 1 to 120977)
Chen, R., Shim, P., Brooks, S., Buehler, E., Chao, Q.,
Johnson-Hopson, C., Khan, S., Kim, C., Altfeld, H., Bel, O.,
Chou, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N.,
Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A.,
Liu, S., Mukharbek, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M.,
Yu, G., Davis, R., Federici, N., Theologis, A. and Ecker, J.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
AUTHORS Department of Biology, University of Pennsylvania, 38th and
REFERENCE Hamilton Walk, Philadelphia, PA 19104-6018, USA
5 (bases 1 to 120977)
Chao, Q., Brooks, S., Buehler, E., Johnson-Hopson, C., Khan, S., Kim, C.,
Shim, P., Altfeld, H., Bel, O., Chen, C., Chou, J., Choi, E., Conn, L.,
Conway, A., Gonzalez, A., Hansen, N., Howing, B., Koo, T., Lam, B.,
Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharbek, N.,
Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A.,
Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Davis, R.,
Federici, N., Theologis, A. and Ecker, J.
TITLE Direct Submission
JOURNAL Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center,
AUTHORS Department of Biology, University of Pennsylvania, 38th and
REFERENCE Hamilton Walk, Philadelphia, PA 19104-6018, USA
On Apr 22, 2000 this sequence version replaced gi:7543634.
COMMENT location/Qualifiers
FEATURES
source

```

CDS

```

1..120977
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F18014"
90..2159
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.1"
/protein_id="AAF79419.1"
/db_xref="GI:8778411"
/translation="MISFEVSPVFTCSIVVNAYCRSGNVDMKAMVFAKETSSGLLE
NVVYNSLNGVAMIGVEGATRVLRMSRGVRNVVYVSLKGYCKKGLMEAEH
VPELLEKRLVADHOWVGVMLMDGCRPGQIRDAVRVDMIEIVRTTTCNSLNG
YCKSGQVLEAQLFSPRNDMSLPDHTTTLTVNGYCRAGVDEALKLCIDMCKQEV
PVTWTVLILKGRIGAFHDVLSLWMKMKRGNADEISCTLEALFKGDFNEAH
KLWENYVARGLDITTLNVMISGLCKMEKVENKEKILDNVNIIFRCRAVQYQALSH
GYKVGILKEAPVKEYMERKGIPTIEMNTLISGAFKYLKHLKADLVLELRGL
TPVATVGAALITGKCNIGMDKAVATCEMIKEKITLVNVCISKIANSLEPLDKDEA
CLIDKTVDFDILLPGYSLKEPFLASATTCCKYKKAESNENSTPKKLYPNTLVN
VALIGLCKAGKEDARKLFSDUSLSDRFIDETVYTLTHGCAAGIDINKAFYLRDEM
ALVGLIPNTVYVALIKGLCKLGVNDVDAQRLHLPLPKGLTPNATVNTTLIDGLVKS
NVAEMKLKEKMEKGLVRSDDKQGVITPREVVLDEPVKLSGTGVIEMNSNEIDVR
RVEENAV"
complement(4647..5702)
/note="putative glycosyltransferase-like protein
emb|CA42905.1; similar to EST gb|A1998490.1"
/codon_start=1
/evidence=not experimental
/product="F18014.2"
/protein_id="AAF79456.1"
/db_xref="GI:8778448"
/translation="MSQHLILILSLHLHKPISATTTIOKFEAPQFVNSADCPLI
DSESDDVVAKPIFCRRRAVHAMTLDAAYIRGSVAIVSLVSHSSCPENIVFPVA
SAGADASLRLATISSFPYLDFTVYVENSVSSVSLISSIRSALDCLNVARSYLADI

```



```

QY 90 ----- 90
Db 22513 CCGAATTAGCTTACTTCACTTGTGTAATATTAGATCTCTCTTACCTTTGATT 22572
QY 91 ---Cys-----
Db 22573 GATTGTACATGATGATTTTGTGTAATGTTATGAGCAACAGGACACAGCCTCTA 22632
QY 95 ProGluIleuIleuAglYthSerSerArgValThrProTyYrSerSerGlnAerGlnSer 114
Db 22633 CCGTGTACATGATGATGATTTTGTGTAATGTTATGAGCAACAGGACACAGCCTCTA 22692
QY 115 ProLeuSerSerAlaPheGlnSerProIleProSerTyYrGlnValSerProSerSer 134
Db 22693 CCGTCTTCTTCCACTTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 22752
QY 135 SerPheProSerProSerArg---GlyGluProAerAerAerAerAerAerAerAer 153
Db 22753 TCATTCCGAGTCTCTTCCGAGTGTGATGATCACAAT-----ATCTCCACAATCTTC 22806
QY 154 ProPheLeuArgAerGlyGlyIleProSerSerIleuProSerLeuAglIleSerAerSer 173
Db 22807 CCGTCTTCTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22866
QY 174 CysProValThrProProValSerSerProThrSerIleuAerProIleProAerPro 193
Db 22867 GCTCCTGCTCTCACCACAGTGTATCCCACTTCTTCCAGATGATGATGATGATGAT 22926
QY 194 TrpGluSerIleuAglYglnSerMetAlaIle---AlaYglnSerMetAlaSerPhe 212
Db 22927 TGGGAATCTTTTACCAACAATCATGATGATGATGATGATGATGATGATGATGAT 22986
QY 213 AsnTyProPheTyYrAlaValSerAlaProAlaSerProThrIleAerGlnPheHis 232
Db 22987 AACTTACCCGTTTATGCGGTGTGACCTGCCGCTCTTCTTCTTCTTCTTCTTCTTCT 23046
QY 233 ThrLeuAlaThrIleProGluCysAerGluSerAerSerSerIleuAerSerGlyHis 252
Db 23047 GCTCCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 23106
QY 253 TrpIleSerPheGlnIlePheAglGlnGlnIleProPheSerAlaSerMetValProThr 272
Db 23107 TGGATTAAGCTTTCAAAAGTTTGACACACACAGCATTCTCTGCTTATGATGATGAT 23166
QY 273 SerProThrPheAerLeuValIleProAlaProGlnIleMetSerProAerThrAlaIle 292
Db 23167 TCCCTTACCTTCAATCTGTAAGACCTGACACACAGCAATGATGATGATGATGAT 23226
QY 293 PheGlnGlnIleGlyGlnSerSerGluPheIlePheGlnAerSerGlnValIleProTyr 312
Db 23227 ATCCAAAGATTTGGTCAAAAGCTCCGAGTTTAAAGTTTAAAGCAACCAAGTTAAG 23286
QY 313 GluGluGluArgIleHisAerValIleMetGluAerLeuGluLeuThrLeuGluAerGly 332
Db 23287 GAAGGAGAGAGATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23346
QY 333 LysAla 334
Db 23347 AAGGCT 23352

```

```

RESULT 10
AF395901 1494 bp mRNA linear PLN 08-JUN-2001
LOCUS AF395901 Lycopodium esculentum mature anther-specific protein LAT61 mRNA,
DEFINITION complete cds.
ACCESSION AF395901 GI:14626760
VERSION AF395901.1 GI:14626760
KEYWORDS
ORGANISM Lycopodium esculentum (tomato)
SOURCE Lycopodium esculentum
Bukavore; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Lamiales; Solanales; Solanaceae; Solanum; Lycopodium.

```

```

REFERENCE 1 (bases 1 to 1494)
AUTHORS McCormick, S.
TITLE Tomato mature anther specific protein (LAT61).
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1494)
AUTHORS McCormick, S.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2001) Plant Gene Expression Center,
USDA/ARS/UC-Berkeley, 800 Buchanan St., Albany, CA 94710, USA
FEATURES
source
1..1494
/mol_type="mRNA"
/db_xref="taxon:4081"
21..1022
/notes="similar to protein encoded by GenBank accession
number AC025808"
/codon_start=1
/product="mature anther-specific protein LAT61"
/protein_id="AAK71662.1"
/db_xref="GI:14626761"
/translation="MMWEAGSPASSAGAGAGGAGVGLPESGGGGGRKPSW
REERNRRRRRAVAAYAKIYTGIRAGCNVNLPGHCONNEVLKALCTEAGMIVEDGT
TYRKCKPMPMEIGTSINTTPSSSRHPSPPSYFASPIISYOSPTSSPSRAD
AMNLSRPSYSLQNVVPSLSLEPLRISNAAPVTPPSSETRHPKOTFNELTKESFAL
NIFEFASAPASPTRVQRTPTIPECHESDSTIDSGOMINFQKVASNVPSPTFNL
VKEPQPLRNDMTIDKXSIDPDPFNVSYKAMEGERIHVGFDLELTGSGNARI"
BASE COUNT 428 a 329 c 302 g 435 t
ORIGIN
Alignment Scores:
Pred. No.: 9..1e-41 Length: 1494
Score: 1060.00 Matches: 219
Percent Similarity: 72.54% Conservative: 32
Best Local Similarity: 63.29% Mismatches: 61
Query Match: 60.06% Indels: 34
DB: Gaps: 8
US-09-995-938a-7 (1-336) x AF395901 (1-1494)
QY 4 AspGluValaThrSerThrSerAlaIleAlaIleAlaIleAlaIleAlaIle----- 21
Db 39 GAATCACACAGCATCTTCTTCCGCGGTCCGAGCTGTGGAATGAGAGTCCGAGATT 98
QY 22 -----ArGArgLysProSerTyrArgGlu 29
Db 99 GGTTTACCGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 158
QY 30 ArgGluAerAerArgArgArgGluArgArgArgArgArgArgArgArgArgArgArgArg 49
Db 159 AGAGAGATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 218
QY 50 GlyLeuAerGluAglGlnIleAerTyYrAerLeuProIleHisCysAerAerAerGluValLeu 69
Db 219 GGTTTAAGAGCTCAAGGAAACTATTAATCTTCCGAGCACTGTGATTAACAATGAAGTCTT 278
QY 70 LysAlaLeuCyValGluAglGlyTyrValIleGluIleAerGlyThrThrTyrArgLys 89
Db 279 AAGGCTTTGTGTAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 338
QY 90 GlyCysLysProLeuProGlyGluIleAglYthSerSerArgValThrProTyYrSer 109
Db 339 GGAATGCAAGCAACCCCGAGAGATGAGAGCACTTCAACAACATCACGCCAAGTTCT 398
QY 110 SerGlnAerGlnSerProLeuSerSerAlaPheGlnSerProIleProSerTyYrGlnVal 129
Db 399 TCAGGATCCAGTCCCTCCATCATCACTTGTGTAAGCCCAATTTCAATTTATCAGCA 458
QY 130 SerProSerSerSerPheProSerProSerArgGlyGluProAerAerAerMetSer 149
Db 459 AGTCCAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 512
QY 150 SerThrPhePheProPheLeuArgAerGlyIleProSerSerIleuProSerLeuArg 169

```

```

Db 513 TCACATCCATATCTTCTCCAAAT---GTGCTTCTCATCTCTTCTCATTAACA 569
Oy 170 IleseraSenSerCyProValThProProValSerSerProThSerThSerValPolys 189
Db 570 AATCAAAACAGAGCCCTTCTTAATCCATCTTCTTCACTCAACT---AGCATCTTAAG 626
Oy 190 ProLeuProAsnTrpGlnSerLeuAlaValGlnSerMetAlaIleAlaValGlnSerMet 209
Db 627 CAAATCTTCAATTTGAAACT-----TTGGCCAAAGAAATCAAG 665
Oy 210 AlaSerPheAsnTrpProPheTrpAlaValSerAlaProAlaSerProThTrhIAsnHis 229
Db 666 TTTCCTTAATAACATCCCTTCTTGTGCTTCAAGCCCAAGCAAGCCCACTAGGCTCAG 725
Oy 220 GlnPheHisThLeuAlaThrlleProGluCyValSerGlnSerAspSerSerThValAlaP 249
Db 726 CATT---ACTCTTCAACTATACCCGAGTGTGAATGATGATCATCTCATCTCATTTGAT 782
Oy 250 SerGlnHisTrpIleSerPheGlnSerPheAlaGlnGlnGlnProPheSerAlaSerMet 269
Db 783 TCAGCCAGCTGATCACTTCAAAAGTAT-----CGCTCAAT 821
Oy 270 ValProTrhSerProThPheAsnLeuValIysProAlaProGlnGlnMetSerProAsn 289
Db 822 GTTCAACCTTCTCCAACTTAACTTGTAAACCTGCTCAGCCCTCTCTCAT 881
Oy 290 ThrAlaAlaPheGlnGluIleGlnSerSerGlnPheLysPheGlnAlaSerGlnVal 309
Db 882 GAT---ATGATCACAGACAAAGGTAAGACATAGACTTGCATTTCAAAATGATCAGTC 938
Oy 310 LysProTrpGlnGluAlaGlnIleHisAspValGlyMetGlnAlaPheGlnLeuThrlleu 329
Db 939 AAGGATGGGAAGGGAAGGATGATCAGATGATGATGATGATGATGATGATGATGATGAT 988
Oy 330 GlyAlaGlnIleValAlaArg 335
Db 999 GGAAGTGGCAATCTCTCC 1016

```

```

RESULT 11
AF372937 945 bp mRNA linear PLN 23-MAY-2001
LOCUS Arabidopsis thaliana Atg19350/F18014_4 mRNA, complete cds.
DEFINITION AF372937.1 GI:13937166
VERSION AF372937.1 GI:13937166
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 945)
Shum,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M.,
Nguyen,M., Seki,M., Southwick,A., Toriumi,M., Yamada,K., Yu,G.,
Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 945)
Shum,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M.,
Nguyen,M., Seki,M., Southwick,A., Toriumi,M., Yamada,K., Yu,G.,
Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE
JOURNAL
REFERENCE
AUTHORS

```

COMMENT

USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and annotation of RAFL cDNAs (RAFL cDNA : RIKEN
Arabidopsis Full-Length cDNA) : Seki,M., Narusaka,M., Ishida,J.,
Saitou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shum,P., Chen,H.,
Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C., Tracy,S.E., Banh,J.,
Bower,L., Chung,M.K., Goldsmith,A.D., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Palm,C.J., Pham,P.K., Quach,H.L., Sakano,H., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yu,G., Davis,R.W., Theologis,A.,
and Ecker,J.R.

Shum,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES

```

source
location/Qualifiers
1..945
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF109-13-019(R12325)"
/notice="ecotype: Columbia"
1..288
289..699
/notice="unknown protein"
/codon_start=1
/product="Atg19350/F18014_4"
/protein_id="AAK50077.1"
/db_xref="GI:13937167"
/translation="MSNAKOSMTSLNPFYAVASAPSTHHROHAPATIECEDSD
SSTVDSGHMISFOKRAOQPPSASVSPSPFNILVKPAPQSPPTAAI1QEIQSGSEF
KFENSOVKPMGERHVDVAMDELTLNGLRHS"
700..945
BASE COUNT 245 a 232 c 172 g 296 t
ORIGIN
3'UTR

```

Alignment Scores:

```

Pred. No.: 71,79e-40 Length: 945
Score: 1035.00 Matches: 202
Percent Similarity: 90.04% Conservative: 6
Best Local Similarity: 87.45% Mismatches: 19
Query Match: 59.64% Indels: 4
DB: 8 Gaps: 3

```

US-09-995-938a-7 (1-336) x AF372937 (1-945)

```

Oy 106 ThrProTrpSerSerGlnAsnGlnSerProLeuSerSerAlaPheGlnSerProIlePro 125
Db 4 ACTCTTCTTCTTCTCCATACCAAGTCTCTTCTTCTTCACTTTGATAGCCCATCTTA 63
Oy 126 SerTrpGlnValSerProSerSerSerSerPheProSerProSerArg--GlyGluPro 144
Db 64 TCTTCAAGACGACAGCTTCTCTTCTTCAATCCGAGCTTCTTCAAGTGGTATGATCA 123
Oy 145 AsnAsnAsnMetSerSerThrpPhePheProPheLeuArgAsnGlyGlyIleProSerSer 164
Db 124 CACAAAT-----ATCTCAACAATCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 177
Oy 165 LeuProSerLeuArgIleSerAsnSerCyProValThProProValSerSerProTrh 184
Db 178 CTTCCTTCACTTGAATTCATAACAGTCTCTCTTCTTCACTTCACTTCACTTCACTTCACT 237
Oy 185 SerLysAsnProLysProLeuProAsnTrpGlnSerIleAlaValGlnSerMetAlaIle 204
Db 238 TCTAGAAACCCCAACCACTTGTGCTTGTGGAATCTTTTACCAAAACAATCATCATTCAG 297
Oy 205 ---AlaLysGlnSerMetAlaSerPheAsnTrpProPheTrpAlaValSerAlaProAla 223

```


Db 298 GCTGCTAAACAGTCAATGCTCTTTGAAGTACCCGTTTATGCGGTCTGCACTGCC 357
 Qy 224 SerProThriHiaGhNlglnPheHieThriLeuAlaThriIleProGluCysAspGluSer 243
 Db 358 AGTCTACTCATCATCGCCAGTTCATGCTCGGCTACTCATATCGAATGTGATGATGCT 417
 Qy 244 AspSerSerThriValAspSerGluHieThriIleSerPheGlnIlyPheAlaGlnGln 263
 Db 418 GACTCTTCACTGTTGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 477
 Qy 264 ProPheSerAlaSerMetValProThriSerProThriPheAsnLeuValIlyProAlaPro 283
 Db 478 CCATTTCTGCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
 Qy 284 GlnGlnMetSerProAsnThriAlaAlaPheGlnGlnIlyGlnIlySerSerGluPhe 303
 Db 538 CAGCAATGTTCTCCAAACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 597
 Qy 304 PheGlnIlySerGlnIlyValIlyProThriProGluGluGluGluGluGluGluGlu 323
 Db 598 TTGTAGAACACCCCAAGTTAACCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
 Qy 324 AspleuGluLeuThriLeuGluGluGluGluGluGluGluGluGluGluGluGlu 334
 Db 658 GATCTAAGCTCAGCGCTTGGAATGTTAAAGCT 690
 RESULT 12
 AX653881 897 bp DNA linear PAT 22-MAR-2003
 LOCUS Sequence 3751 from Patent WO03000898.
 DEFINITION AX653881
 ACCESSION AX653881
 VERSION AX653881.1 GI:29156695
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa
 Oryza sativa
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriocaulaceae; Oryzaceae; Oryza.
 REFERENCE
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
 Katagiri, F., Quan, Y., Tao, Y., Whitman, S., Xie, Z., Zhu, T. and Zou, G.
 TITLE Plant genes involved in defense against pathogens
 JOURNAL Patent: WO 0300898-A 3751 03-JAN-2003;
 Syngenta Participation AG (CH)
 FEATURES
 SOURCE 1..897
 /organism="Oryza sativa"
 /mol_type="genomic DNA"
 /db_xref="taxon:4530"
 BASE COUNT 147 a 328 c 309 g 112 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,32e-32 Length: 897
 Score: 872.50 Matches: 184
 Percent Similarity: 68.92% Conservative: 40
 Best Local Similarity: 56.62% Mismatches: 62
 Query Match: 49.43% Indels: 39
 DB: 6 Gaps: 10
 US-09-995-938A-7 (1-336) x AX653881 (1-897)
 Qy 16 AlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 35
 Db 7 TCCGG 66
 Qy 36 ArgAlaArgArgArgArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 55
 Db 67 CGGAGGG 126
 Qy 56 AspGlyThrAsnLeuProGluGluGluGluGluGluGluGluGluGluGluGluGlu 75
 Db 127 AACCTAACACCTCCCGACAGCTGCGACAAACAGAGGTGCTCAAGGCCCTCTGCGCGAG 186

Qy 76 AlaGlyTrpValValGluGluAspGluThriThriThriThriThriThriThriThriThri 95
 Db 187 GCCGGCTGGGTTGTCAGAGACGACGACGACGACGACGACGACGACGACGACGACGACG 246
 Qy 96 GlyIleIleAlaGlyThriSerSerArgValThriProTyrsSerSer---GlnAsnGlnSer 114
 Db 247 TCGTGGCTGGGGAGCGCTGGTGGGAGTGAAGCCCTGCTGTCACAGCAGCTGTGAGC 306
 Qy 115 ProLeuSerSerAlaPheGlnSerProIleProSerThriThriThriThriThriThri 134
 Db 307 GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
 Qy 135 SerPheProSerProSerArgGlyIleProAsnAlaAsnMetSerSerThriPhePhePro 154
 Db 367 AGCTTCCGAGCCGACGCGG-----ATCAGAACCCGAGGCGCTCTGCTCTCCG 420
 Qy 155 PheLeuArgAsnGlyGlyIleProSerSerLeuProSerLeuArgIleSerAsnSerCys 174
 Db 421 TTCTCTCGG---T---GGGCTCCC---AACCTTCGGGCTCTCCGCTCCAGACGCGG 471
 Qy 175 ProValThriProProValSerSerProThriSerThriThriThriThriThriThri 192
 Db 472 CCCGTACGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
 Qy 193 AsnTrpGluSerIleAlaValGlnSerMetAlaIleAlaValGlnSerMetAlaSerPhe 212
 Db 532 GACTGGGAC-----GTCAGCCGCTTC 552
 Qy 213 AsnTrpProPheThriAlaValSerAlaProAlaSerProThriHiaGlnPheHie 232
 Db 553 CGGACCCCTTCTTCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGAG 612
 Qy 233 ThriLeuAlaThriIleProGluGluAspGluSerAsnSerSerThriValAspSerGluHie 252
 Db 613 CACCGGACAGACGATACCGAGATGCGACAGATGCGACGATGCGACGATGCGACGATG 672
 Qy 253 TrpIleSerPheGlnIlyPheAlaGlnGlnIlyProPheSerAlaSerMetValProThri 272
 Db 673 TGATACAGCTTGCGG-----ATGGCCAGCGCGCGCGCGGAG 708
 Qy 273 SerProThriPheAsnLeuValIlyProAlaProGlnGlnMetSerProAsnThriAla 292
 Db 709 TCGGCCACCTTCAACCTCTGTCACCGG-----GGCGCTCTCCACCTTCAAC 753
 Qy 293 PheGlnIlyIle-----GlyIle---SerSerGluPheIlyPheGluAsnSer 307
 Db 754 TCCATGAGATGAAAGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 813
 Qy 308 GlnValIlyProTrpGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 327
 Db 814 AGGCTGACGCCATGAGAGGCGGAGAGATCCACGAGTTCGCGCGCGAGAGAGCTTCAAGCT 873
 Qy 328 ThrLeuGluAsnGly 332
 Db 874 ACGCTCGCGCGTGGC 888
 RESULT 13
 AY050394 1389 bp mRNA linear PLN 20-AUG-2001
 LOCUS Arabidopsis thaliana AT4g36780/C7A10.580 mRNA, complete cds.
 DEFINITION AY050394
 ACCESSION AY050394
 VERSION AY050394.1 GI:15215730
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (chale crese)
 Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosoids; euroside II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 AUTHORS 1 (bases 1 to 1389) Koese, E., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Shin, P.,
 Ban, J., Bower, J., Carninci, P., Dale, J.M., Goldsmith, A.D.,
 Hayashizaki, Y., Ishida, J., Jung, P.X., Jones, T., Kamita, A.,

Db 778 CAACTTCTATGCGCATGATATGAAAGTGCAGCTGGGGAATGTCGGGAATGATGG 837
Qy 298 GlnSerSerGluPhePheGluAenSerGlnValLeuProTTrpGluValGlyLe 317
Db 838 AGAGGGCGAGAGTTGAGTTGAGTAAGTAAGACGTTACCGCGGAGAGCGGAATGAT 897
Qy 318 HisAspValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGlyValValArg 335
Db 898 CATGAAGTTGGCGTACAGATCTCGAACTCCTCTGCGGAGCACTTAAGCCAGA 951

RESULT 15
AP004276 144480 bp DNA linear PLN 12-JUN-2003
LOCUS AP004276
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
ACCESSION AP004276
VERSION AP004276.2 GI:31621040
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0453G03
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 144480)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
JOURNAL Direct Submission
Submitted (17-OCT-2001) Takuji Sasaki, National Institute of
Agricultural Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7466)
On Jun 12, 2003 this sequence version replaced gi:16215797.
COMMENT The orientation of the sequence is from SP6 to T7 of the PAC clone.
FEATURES
Location/Qualifiers
source
1..144480
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="7"
/clone="P0453G03"

BASE COUNT 38130 a 33863 c 32771 g 39716 t
ORIGIN

Alignment Scores:
Pred. No.: 3,866-29 Length: 144480
Score: 856.00 Matches: 187
Percent Similarity: 59.74% Conservative: 40
Best Local Similarity: 49.21% Mismatches: 71
Query Match: 48.50% Indels: 82
DB: 8 Gaps: 11

US-09-995-938a-7 (1-336) x AP004276 (1-144480)

Qy 4 AapGlyAlaThrSerThrSerAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaArg 23
Db 92886 GAGGGAGAGTGAAGGAGAGTGTGTATGAGCTCCGGGCGGCGGCGGAGG 92945
Qy 24 LysPProSerTrpArgGluArgGluAenArgArgGluArgGluArgGluArg 43
Db 92946 ACCCGAGCTGGAAGAGAGGAGAAACAAGAGCGCGGAGCGCGGCTGCCATC 93005
Qy 44 AlaAlaValGlyLeuThrGlyLeuArgGluGlnGlyAspTrpAsnLeuProLysHisCys 63
Db 93006 GCCGCAAGATCTTCAACGGGCGCTCCGGGAGTAACCTCCCAAGCACTGC 93065
Qy 64 AspAsnAsnGluValLeuValLeuValLeuValGluValGluValGluValGlu 83

Db 93066 GACAAACAGAGAGTCTCTCAAGCGCTCTCCCGCCAGCGCGGCTTCCAGAGACAC 93125
Qy 84 GlyThrThrTyxArgLys----- 89
Db 93126 GGCACACCTACACGAGAGGTCAATTCAATCTTCCCGTCTCATGTCGTGCTGTT 93185
Qy 89 ----- 89
Db 93186 GCTGTCAATTCTTGATCTTGATGTAGATCTGAGCCGAGCTGAGCTGCTGT 93245
Qy 90 -----GlyCysLeuProLeuProGlyLeuValLeuGly 100
Db 93246 TTCTTGTTTCTTTTGTGTCGTCGAGAGATTAAGCCGCCCATCTGCGCTGGGGA 93305
Qy 101 ThrSerSerArgValThrProTyxSerSerGlnAsnGlnSerProLeuSerAla 119
Db 93306 GCGTCGAGGAGATGAGCCCTCTGTCACAGCAGCTGAGCGCGCGCTGCTGTG 93365
Qy 120 PheGlnSerProLysProSerTrpGlnValSerProSerSerSerPheProSerPro 139
Db 93366 TTCGAGCCCGGTGCTGTCGACCAAGCAGCCGCGGCTGTCGAGCTTCCGAGCCCC 93425
Qy 140 SerArgGlyGluProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAsnGly 159
Db 93426 AGCCGG-----ATGACAAACCGAGCGCTCTGCTCTCTCTCTCTCTCTCTCT 93473
Qy 160 GlyLeuProSerSerLeuProSerLeuArgGlnSerAsnSerCysArgValThrProPro 179
Db 93474 GGGCTCCCC---AACCTCCCGCGCTCCGCGCTCCAGAGCGCGCTCCACCGCGCGG 93530
Qy 180 ValSerSerProThrSerLeuAsnProLysProLys-----ProAsnTrpGluSerIle 197
Db 93531 CTCTGTCGCGGACGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 93584
Qy 198 AlaLysGlnSerMetAlaIleAlaLysGlnSerMetAlaSerPheAsnTrpProPheTy 217
Db 93585 -----GTCGACCCCTTCCGAGCCCTTCTTC 93611
Qy 218 AlaValSerAlaProAlaSerProThrAlaArgGlnGlnPheAlaThrIle 237
Db 93612 GCGGCTCCGG 93671
Qy 238 ProGluCysAspGluSerAspSerSerThrValAspSerGlyHisTrpLysSerPheGln 257
Db 93672 CCGGAGTCGAGCGAGTCCGAGCTCTCCAGCGGAGCTCCGCGCGGAGTCAAGCTTCCAG 93731
Qy 258 LysPheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsn 277
Db 93732 -----ATGGCACAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 93767
Qy 278 LeuValLysProAlaProGlnGlnMetSerProAsnThrAlaAlaPheGlnGluLe 296
Db 93768 CTCGTCAACCG-----GGCGCTCACTCCCACTCCATGAGATAGAA 93812
Qy 297 -----GlyGln--SerSerGluPheLysPheGluAenSerGlnValLysProTrp 312
Db 93813 GCGACG 93872
Qy 313 GlnGluGluArgGlnLeuAspValGlyMetGluAspLeuGluLeuThrLeuValAsnGly 332
Db 93873 GAGGGCGAGAGATCTCAAGAGTCCCGCGCGAGCTCGAGCTCAAGCTCCGCGCTGCG 93932

Search completed: December 20, 2003, 05:04:02
Job time : 4467 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 20, 2003, 01:22:20 ; Search time 321 Seconds

(without alignments)
2825.580 Million cell updates/sec

Title: US-09-995-938a-7

Perfect score: 1765
Sequence: 1 MTSDGATSTSAAAAAAAAA.....IHVGNEDLELTGNGKARG 336

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h
-O=/cgn2.1/US09995938/rnatat.19122003.142606.4227/app_query_faasta_1.519
-DB=N_Geneseq.19Jun03 -OPMT=faastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonmb2 -TRANS=human40.cd1
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09995938 @CGN 1.1 312 @rnatat.19122003.142606.4227 -NCPUS=6 -ICPU=3
-NO MAP -LARECOVERY -NEG_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N_Geneseq.19Jun03.:

```
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*
```

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	1758	99.6	1501	21	AA37477
2	1511.5	85.6	1359	21	AA37137
3	1511.5	85.6	1418	21	AA36495
4	825	46.7	881	24	AB212580
5	541	30.7	1213	21	AA51043
6	541	30.7	1217	21	AA42550
7	496.5	28.1	642	21	AA42317
8	371	21.4	459	22	AA87878
9	371	21.0	459	25	ABX22943
10	321.5	18.2	284	25	ABX32572
11	171.5	9.7	1050	22	AA21685
12	171.5	9.7	2849	22	AA21684
13	159.5	9.0	2800	22	AA207131
14	156.5	8.9	2805	22	AA207130
15	153.5	8.7	227968	24	ABK83497
16	149.5	8.5	2803	22	AA207132
17	148.5	8.4	2586	22	AA21064
18	147.5	8.4	2582	22	AA21062
19	147	8.3	156	25	ABX75052
20	146.5	8.3	6521	23	ABX8121
21	143.5	8.1	31422	21	ABX11153
22	143.5	8.1	31422	22	AA92302
23	141.5	8.0	12001	16	AA76213
24	141	8.0	6286	22	AA75248
25	140	7.9	9408	21	AA470157
26	140	7.9	9636	15	AA67190
27	138.5	7.8	1428	24	ABO91674
28	138	7.8	1313	22	AA74867
29	137.5	7.8	3182	10	AA24982
30	137.5	7.8	3182	20	AA64276
31	137.5	7.8	3182	21	AA474385
32	137.5	7.8	3182	21	AA424123
33	137.5	7.8	3182	21	AA236712
34	137.5	7.8	3182	22	AA514694
35	137.5	7.8	3182	22	AA68402
36	137.5	7.8	3182	22	AA78512
37	137.5	7.8	3182	24	ABK84562
38	137.5	7.8	3182	24	ABN97226
39	137.5	7.8	3182	24	ABK64811
40	137.5	7.8	3182	24	ABK38313
41	137.5	7.8	3182	24	AB62476
42	137.5	7.8	3182	24	AB688572
43	137.5	7.8	3182	25	ACA10642
44	137.5	7.8	3182	25	ABX99593
45	137.5	7.8	3182	25	ABX10949

ALIGNMENTS

RESULT 1
ID A37477
AAC37477 standard; DNA, 1501 BP.
AC A37477:
AC A37477:
17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17521.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.

XX 06-SEP-2000.
PD XX
PF 25-FEB-2000, 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0133548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137722.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0140997.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147320.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0155659.
PR	29-SEP-1999;	99US-0155696.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Alignment Scores:	
Pred. No.:	7,09e-102
Score:	1758.00
Percent Similarity:	99.70%
Best Local Similarity:	99.70%
Query Match:	99.60%
DB:	21
US-09-995-938A-7 (1-336) x AAC37477 (1-1501)	
	Length: 1501
	Matches: 335
	Conservative: 0
	Mismatches: 1
	Indels: 0
	Gaps: 0

Qy	1	MetThrSerArgGlyValAlaHisSerThrSerAlaAlaAlaAlaAlaAlaAla	20
Db	187	ATGACTTCGGATCGAGACTACCTCGACATCAGCAGCTCGACCTCGCGCGCGCAGCAGCG	246
Qy	21	AlaArgArgGlyPProSerTyrArgGluArgGluAsnAsnArgArgArgGluArgArg	40
Db	247	GCGAGAGAGAACCCGCTCGGAGAGAAAGAGAAATATTCGGAGAGAGAAACGGAGA	306
Qy	41	ArgAlaValAlaAlaLysIleTyrThrGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro	60
Db	307	AGAGCTGTAGCTCGCAGACATATACATCGGCTTAGAGCTCAAGTGTATTAATTTGGCT	366
Qy	61	LysHisGlyAspAsnAsnGluValLeuLysValLeuCybValGluAlaGlyTyrPvalVal	80
Db	367	AAACCTTGGAATATATAGAGCTTAAAGCTCTTGCTTGAAGCTGGTTGGCTGTT	426
Qy	81	GluGluAspGlyThrThyTyrArgLysGlyCybLysPProLeuProGlyGluIleAlaGly	100
Db	427	GAGAGAGAGGCTATCTTATCTTATCCAGAGGATCGAACCTTTACTGCTGATGATGCTGGG	486
Qy	101	ThrSerSerArgValThrProTyrSerSerGlnAsnGlnSerProLeuSerSerAlaPhe	120
Db	487	ACTTCATCTCGAGTAATCTCATTTATATACAGAAACAGAGCCCTCTTTATCAGCCTTT	546

QY	121	GlnSerProIleProSerTYrGlnValSerProSerSerSerPheProSer	140
Db	547	CAAGTGTCCATGCCATCTTACCAAGTTAAGCCGCTTCTTCATCATTC	606
QY	141	ArgGlyIuProhAnAenMetSerSerThPhehPheProPheLeuArgAsnGly	160
Db	607	CGGGTACCAACCAATACCAACATGTCTCTACATCTTCTCCCTTCTC	666
QY	161	IleProSerSerIleuProSerLeuArgIleSerAsnSerCysProValThrProProVal	180
Db	667	ATTCTCTCTCTCTCTCTCTCTCCAGATCTCAACACGTTGTCCAGTTACCC	726
QY	181	SerSerProThSerIleValAsnProLysProLeuProAsnTrpGluSerIleAlaLysGln	200
Db	727	TCAATGCCACCTTTAGAACCCGGAAACCTTTGCTTAACTGGGAATTAATGCTTAAGCA	786
QY	201	SerMetAlaIleAlaLysGlnSerMetLysSerPheAsnTYrProPheTYrAlaValSer	220
Db	787	TTCATGGCCATATGTTAAACATCAATGAGCGCTTATTAATCTTCTCATGCGGTTCT	846
QY	221	AlaProLysSerProThrIleAlaGhIleGlnPheIleThrLeuAlaThrIleProGluCys	240
Db	847	GCACCTCTGCTGTCGACACATCGGCACAGTTTCATACCCCGCTACTATACCTGAATGT	906
QY	241	AspGluSerAspSerSerSerThValAspSerGlyNHistripIleSerPheGlnLysPheAla	260
Db	907	GATAGTCTGACTCTTCCACTGTGTTGATTCGGTCAATGGATTAAGCTTTCAGAACTTGCA	966
QY	261	GlnGlnIleProPheSerAlaSerMetValProThrSerProThrPheAsnLeuValLys	280
Db	967	CAACAAACAGCACTCTCTGCTCTATGTCACAACTCTCTACTTCAATCTTGTA	1026
QY	281	ProLysProGlnIleMetSerProAsnThrAlaAlaPheGlnIleGlyGlnSerSer	300
Db	1027	CTGTGCGCTCAGCAGATGCTCCAAATACGTGCTTCCAAAGATTTGCAAAAGCTCT	1086
QY	301	GluPheLysPheGluAsnSerGlnValLysProTrpGluGluArgIleHisAspVal	320
Db	1087	GAGTTTAAATTGTGAGATACCAAGTTAAACCCCTGGGAAGAGAGAGATCATGATGTG	1146
QY	321	GlyMetGluAspLeuGlnLeuThrLeuGlyAsnGlyLysAlaArgGly	336
Db	1147	GGTATGAGAGATCTTGAGCTTACCTTGAATGGAAAGGAGGCTCGTGT	1194
RESULT 2			
AACG7137			
ID	AACG7137	standard; DNA; 1359 BP.	
AC	AACG7137;		
XX	17-OCT-2000	(first entry)	
DT	17-OCT-2000		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 16310.		
XX	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway;		
KW	metabolic pathway; promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
XX	EP1033405-A2.		
PN	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
PF	25-FEB-1999; 99US-0121825.		
XX	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	29-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128214.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141267.
PR 01-JUL-1999; 99US-0141642.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149375.
PR 17-AUG-1999; 99US-0149178.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161921.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 2,08e-86 Length: 1359
 Score: 1511.50 Matches: 296
 Percent Similarity: 91.07% Conservative: 10
 Best Local Similarity: 88.10% Mismatches: 25
 Query Match: 85.64% Indels: 5
 DB: 21 Gaps: 4

US-09-995-938a-7 (1-336) x AAC37137 (1-1359)

Qy 1 MetThSerAePglYAlaThSerTherSerAlaAlaAlaAlaAlaAlaAla 20
 Db 112 ATGACGCTTGAACGAGCAACGTCACGTC---GCTGACGCTGACGACGACGACG 168
 Qy 21 AlaArgArgLysProSerTrpArgGluArgGluAenAenArgArgArgArgArg 40
 Db 169 ACGAGAGCAACCGCTCGGAGAGAGGAGAACATCGAGAGAGAGAGAGAGAGAG 228
 Qy 41 ArgAlaValAlaAlaLysIleTyrThrGlyLeuArgAlaGlnGlyAspTyrAenLeuPro 60
 Db 229 AGACCTGTGCGCGGACAAATTAATCTGCTTACGCTCAAGCTCAAGTAACTCAATCTTCCA 288
 Qy 61 LysHisCysAaPaaAenGluValLeuLysAlaLeuCysValGluAlaGlyTrpValVal 80
 Db 289 AAACATTGTGACAAACAAATGAGGTCTTAAGCTCTTGTCTGAAGCTGGTGGGTGT 348
 Qy 81 GluGluAePglYThrThrTyrArgLysGlyCysLysProLeuProGlyGluIleAlaGly 100
 Db 349 GAAGAAAGCGAAGCTTAATATGCAAGGACAAAGCCCTTACCTGGTGCACATGGCTGA 408
 Qy 101 ThrSerSerArgValThrProTyrSerSerGlnAenGlnSerProLeuSerSerAlaPhe 120
 Db 409 TCATCTTCTCGAGCAATCTTACTTCCATTAACCAAGTCTCTTCTTCTTCTTCTTCTT 468
 Qy 121 GlnSerProIleProSerTyrGlnValSerProSerSerSerPheProSerProSer 140
 Db 469 GATACCCCATCTTAATCTTAACCAAGTCACTCTTCTTCTTCTTCTTCTTCTTCTTCT 528
 Qy 141 Arg---GlyGluProAenAenAenMetSerSerThrPhePheProPheLeuArgAenGly 159
 Db 529 CGAGTTGGTGAATCCACAAAT-----ATCTCCAAATATTTCTTCTTCTTCTTCTTCTT 582
 Qy 160 GlyIleProSerSerLeuProSerLeuArgIleSerAenSerCysProValThrProPro 179

Db 583 GGATTCCTCATCGCTTCTTCCACTTGAATCTCAACAGAGCTCTGTCACTCACCA 642
 Qy 180 ValSerSerProThrSerLysAenProLysProLeuProAenTrpGluSerIleAlaLys 199
 Db 643 GTGTCAATCCCACTTCTTCAAGTCAATCTTCTTCAAGTCAATCTTCTTCTTCTTCTT 702
 Qy 200 GlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPheAenTyrProPheTyrAla 218
 Db 703 CAATCAATGCTCAGCTGCTTAACAGTCAATGATCTTCTTCAACTACCGCTTTATGCG 762
 Qy 219 ValSerAlaProAlaSerProThrHisArgHisGlnPheHisTrpLeuAlaThrIlePro 238
 Db 763 GTGTCTGACCTGCGACATCTTCACTCATGATGCGCATGCGCATGCGCATGCGCATGCT 822
 Qy 239 GluCysAePglYSerAePglYSerSerThrValAePglYHisTrpIleSerPheGlnLys 258
 Db 823 GAATGTATGATGCTGATCTTCCACTTCACTTATTTCTGCTCATGATGATGATGATGAT 882
 Qy 259 PheAlaGlnGlnGlnPhePheSerAlaSerMetValProThrSerProThrPheAenLeu 278
 Db 883 TTTCACACAAACAGCCATCTTCTGCTTATGAGCCAACTCGCTTCACTTCAATCTC 942
 Qy 279 ValLysProAlaProGlnGlnMetSerProAenThrAlaAlaPheGlnGluIleGln 298
 Db 943 GTGAACCTGACCAACAGCAATGCTTCTTCAACACAGAGCAATCCAAAGATGATGAT 1002
 Qy 299 SerSerGluPheLysPheGluAenSerGlnValLysProTrpGluGlyGluArgIleHis 318
 Db 1003 AGCTCCGAGTTTAAATTGAGAACACAGCAATGATGATGATGATGATGATGATGAT 1062
 Qy 319 AspValGlyMetGluAePglYLeuThrLeuGlyAenGlyValAla 334
 Db 1063 GATGTGCTATGAGAGATCTAGAGCTCAAGCTTGAATGATGATGAT 1110
 RESULT 3
 AAC36495
 ID AAC36495 standard; DNA; 1418 BP.
 AC AAC36495;
 XX AC
 DT 17-OCT-2000 (first entry)
 XX DT
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14021.
 XX DE
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX OS
 PN EP1033405-A2.
 XX PN
 PD 06-SEP-2000.
 XX PD
 PF 25-FEB-2000; 2000EP-0301439.
 XX PF
 XX 25-FEB-1999; 99US-0121825.
 XX 05-MAR-1999; 99US-0123180.
 XX 09-MAR-1999; 99US-0123548.
 XX 23-MAR-1999; 99US-0125788.
 XX 25-MAR-1999; 99US-0126264.
 XX 29-MAR-1999; 99US-0126785.
 XX 01-APR-1999; 99US-0127462.
 XX 06-APR-1999; 99US-0128234.
 XX 08-APR-1999; 99US-0128714.
 XX 16-APR-1999; 99US-0128845.
 XX 19-APR-1999; 99US-0130077.
 XX 21-APR-1999; 99US-0130449.
 XX 23-APR-1999; 99US-0130510.
 XX 28-APR-1999; 99US-0130891.
 XX 30-APR-1999; 99US-0131449.
 XX 30-APR-1999; 99US-0132048.
 XX 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132488.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137422.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140355.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142555.
PR 06-JUL-1999; 99US-0142990.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147703.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139819.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142927.
 PR 13-JUL-1999; 99US-0143547.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 19-JUL-1999; 99US-0144336.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145224.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152353.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155119.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158359.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 22-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
 Pred. No.: 2,24e-25
 Score: 541.00
 Percent Similarity: 48.31%
 Best Local Similarity: 40.17%
 Query Match: 30.65%
 DB: 21

Length: 1213
 Matches: 143
 Conservative: 29
 Mismatches: 64
 Indels: 120
 Gaps: 17

US-09-995-938a-7 (1-336) x AAC51043 (1-1213)

```
OY 23 ArgLysProSerTyrArgGluArgGluAsnAsnArgArgArgGluArgArgArgAla 42
Db 163 AGAAGCGCCGACTGGAAAAGAGAGAGAACAAACGCGAGAGCGCGGAGAGACGAGCG 222
OY 43 ValAlaAlaValIleTyrThrGlyLeuAlaGlnGlyAspTyrAsnLeuProLysHis 62
Db 223 ATGGCGGCTAGATCTTCGAGAGACTTACGAGATTCATGAACTTCAGGCTCCCTAAAC 282
OY 63 CysAspAsnAsnGluValLeuLysAlaLeuCysValAlaAlaGlyTyrValValGluGlu 82
Db 283 TCGGCAACAAATGAAAGTCTCTCAAGCTTTATGATGAAAGCGTTGAGCTGAGAACGC 342
OY 83 AspGlyThrThrTyrArgGlyGlyCysLysProLeuProGlyGluIleAlaGlyThrSer 102
Db 343 GACGGAACTACTTACCGCAAGGAGTGCACAAACCAATG----- 378
OY 103 SerArgValThrProTyrSer---SerGlnAsnGlnSerProLeuSerSerAlaPheGln 121
Db 379 GATCGAATGAGACTCATGATGATGTTCTACTCTAGCTAGTCCATGCTCATCGTATCAACAT 438
OY 122 SerProIleProSerTyrGlnValSerProSerSerSerSerPheProSerProSerArg 141
Db 439 AGCCCTCGTCTTCTTCAATCAAGCCCTTCTGCTTCATCATCTCCGAGTCTTCAACAC 498
OY 142 ----GlyGluProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAsn--- 158
Db 499 CCATTGGATGATGCATAC-----TCACATAATCCCATGCTGCTCAAGAACCTCC 543
OY 159 ---GlyGlyIleProSerSerLeuProSerLeuArg---IleSerAsnSerCysProVal 176
Db 544 TCTTCAAACTCACTTCAAGCTTCCCTTCTTCAATGAAATTCATATAGACCTCCCGT 603
OY 177 ThrProValSerSerSerProThrSerLysAsnProLysProLeuProAsnTrpGluSer 196
Db 604 ACTCCGCAATGGCT----- 618
OY 197 IleAlaLysGlnSerMetAlaIleAlaLysGlnSerMetAlaSerPheAsnTyrProPhe 216
Db 618 ----- 618
OY 217 TyrAlaValSerAlaProAlaSerProThrIleArgHisGlnPheHisThrLeuAlaThr 236
Db 619 -----CGAAGCCCTACT---CGATATCAA-----GTAACC 645
OY 237 IleProGluCysAspGluSerAspSerSerSerThrValAspSerGlyHisTrpIleSerPhe 256
Db 646 ATCCCT-----GACTTCTGA---TGGCTCTCA--- 669
OY 257 GlnLysPheAlaGlnGlnInProPheSerAlaSerMetValProThrSerProThrPhe 276
Db 670 ----GGATGCAAACTCCGACAGCGCA-----CCGCTTCTCTACTTTC 711
OY 277 AsnLeuVal----- 279
Db 712 AGTTAGTTTTCAGAAACCCGTTTTCGACAAAGAGCTTTAAATGGAGATTTGTAAT 771
OY 280 -----LysProAla---ProGlnGlnMet 286
Db 772 TCACCAATGTGACTCTCGGACAAAGTGAAACCTCTCCAGCTATTCCTGCTGCTT 831
OY 287 SerProAsnThrAlaAlaPheGlnGlnIleGlyGlnSerSerGlnPheLysPhe----- 304
Db 832 GATCAAGACTCTGATGTGCAATGCGCTAGTGAAGACGCGTGAAGTTCCGTTTGCTTGT 891
OY 305 -----GluAsnSerGlnValLysProTrpGlnGlyGluArgGlnHis---Asp 319
Db 892 AACGCAATGCTGCAATGGAATGGAAGCTTGGGAAAGGAGAAAGATACATGAGACAA 951
OY 320 ValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGlyLysAlaArg 335
```

```
Db 952 TGTGTTGAGATGATTGAACTTACACTTGGAACACTCAAGAACCGAGA 999
RESULT 6
AAC42550
ID AAC42550 standard; DNA; 1217 BP.
XX
XX AAC42550;
AC
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35978.
DE
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35978.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
```

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 22-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142155.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144881.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147036.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154019.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159741.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 2,256-25
Score: 541.00
Percent Similarity: 48.31%
Best Local Similarity: 40.17%
Query Match: 30.65%
DB: 21
Gaps: 17

US-09-995-938a-7 (1-336) x AAC42550 (1-1217)

OY 23 ArgLysProSerTyrArgGluuArgLysuAsnArgArgArgGluuArgArgArga1a 42
DB 163 AGACGCCGACTTGGAAAGAGAGAGAAACAAACGCGGAGCGCGCAAGACGAGCC 222

43 ValAlaAlaLeuAlaLeuThyThyGlyLeuAlaGlnGlyAspTyrAsnLeuProLeuHis 62
223 ATGGCGGCTAAGATCTTCAGAGACTAAGATTCAGAAACTTCAGAGCTCCCTAAAC 282
63 CysAspAsnAsnGluValLeuLeuValAlaLeuCysValGluAlaGlyTyrValValGluGlu 82
283 TCCGACAAAGAAATGAACTCTCAAGCTTTATGCAAGAAAGCTGGTGGACTGAGAAAGAC 342
83 AepGlyThrThrThyArgGlyGlyCysLeuProLeuProGlyGluLeuAlaGlyThrSer 102
343 GACGGAACCTTACCGCAAGGATCCAAACCAATG----- 378
103 SerArgValThrProTyrSer---SerGlnAsnGlnSerProLeuSerSerAlaPheGln 121
379 GATCGAATGACCTCATGATGTTCTACTTACGAGTCCATGCTCATGATCATCAACT 438
122 SerProIleProSerTyrGlnValSerProSerSerSerSerPheProSerProSerArg 141
439 AGCCCTCGCTTCTTCAATCCAAACCCCTTCCTTCATCATCCCGAGTCTCAACAC 498
142 -----GlyGluProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgan--- 158
499 CCATTGGTGGATCTAAC-----TCCTAATCCCATGCTCAAGAACCTC 543
159 ---GlyGlyIleProSerSerLeuProSerLeuArg---IleSerAsnSerCysProVal 176
544 TCTTCAAACTCACTTCCAAAGCTTCCCTTCCATCGAAATTCATTAAGCCCTCCCG 603
177 ThrProProValSerSerProThrSerLeuAsnProLeuProLeuProAsnTyrGluSer 196
604 ACTCCGCAATGGCT----- 618
197 IleAlaLeuGlnSerMetAlaIleAlaLeuGlnSerMetAlaSerPheAsnTyrProPhe 216
618 ----- 618
217 TyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThr 236
619 -----CGAAGCCCTACT---CGTGAATCA-----GTAACC 645
237 IleProGluCysAspGluSerAspSerSerThrValAspSerGlyHisTyrIleSerPhe 256
646 ATCCCT-----GACTTCGGA---TGCTCTCA--- 669
257 GlnLeuPheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPhe 276
670 -----GGAATGCAAACTCCGACGACCGA-----CCGCTTCTCTCTACTTC 711
277 AsnLeuVal----- 279
712 AGTTTAGTTTCAAGAAACCGTTTTCACAAAGAGGCTTTAAATGAGGATTTGTAAT 771
280 -----LysProAla---ProGlnGlnMet 286
772 TCACCAATGTGACTCTGACAAAGTGGAACGTGCTCCAGCTATTCCTGCTGTGT 831
287 SerProAsnThrAlaAlaPheGlnGlnIleGlyGlnSerSerGlnPheLysPhe----- 304
832 GATCAGAACTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 891
305 -----GluAsnSerGlnValLysProTyrGluGluAlaGlnLeuHis---Asp 319
892 AACCGAATGCTGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
320 ValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGlyLysValArg 335
952 TGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999

XX 17-OCT-2000 (first entry)
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 35093.
XX
DE Hybridization assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
PN EPI03405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.

PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139753.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140685.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142330.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144086.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144335.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144335.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144684.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145224.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145918.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148665.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158222.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159284.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT	

QY 60 ProLYHhACyAspAenGluValLeuLYsAlaLeuCyValGluValGlyTTPVal 79
 DB 288 CTTAAAGCCTGGATACAAACGAGGTCTTAAGCTCTGCTCGAAGCTGGTGGATC 348
 QY 80 ValGluGluAspGlyThrThrTYrArgLYsGlyCyAlpProLeuProGlyGluLea 99
 DB 349 GTCGAAGAGATGGACCACTTATCCAGAGGGTTAG-CGACCACATCAGATATTCA 407
 QY 100 GlyThrSerSerArgValThrProTYrSerSerGluAenGlnSerProLeuSerSera 119
 DB 408 GGAACCTCCATCAAACTTCAGACAAATTCATCATCAACCAAGTCCACATCATCAGCT 467
 QY 120 PheGlnSerProIleProSerTYrGlnValSerProSerSerSerSerPheProSerPro 139
 DB 468 TTTCAGAGTCTGCACCTTGATACACGAGAGTCCAGTCTCATCATCTCTCCGAGTCCA 527
 QY 140 SerArgGlyGluProAenAenAenMetSerSerThrPhePheProPheLeuArgAen--- 158
 DB 528 TCTCGCTATGAC---CGAAACCTTCTTCTCATCTCTCTCTCTCTCTCTCTCTCTCA 584
 QY 159 ---GlyGlyIleProSerSerLeuProSerLeuArgIleSerAenSerCySProVal 176
 DB 585 GCTTCTCGATTCCTGCTAACCTTCACCTTAGATATCCAAAGTGGCCTGTG 641
 RESULT 8
 AAH87878
 ID AAH87878 standard; cDNA, 459 BP.
 AC AAH87878;
 XX 25-SEP-2001 (first entry)
 DT Peppermint plant oil gland expressed cDNA 234.
 DE Peppermint plant oil gland expressed cDNA 234.
 KW Peppermint; plant oil gland cell; terpenoid essential oil; resin;
 KM genetic mapping; antisense suppression; recombinant expression; ss.
 XX Mentha x piperita.
 OS
 XX WO200153319-A1.
 PN 26-JUL-2001.
 PD 19-JAN-2001; 2001WO-US02567.
 PF 20-JAN-2000; 2000US-0177264.
 PR
 XX (CROT/) CROTEAU R B.
 PA (LANG/) LANGE B M.
 XX (WILD/) WILDUNG M R.
 XX Croteau RB, Lange BM, Wildung MR;
 DR WPI: 2001-488706/53.
 XX New nucleic acid molecules corresponding to mRNA molecules expressed in
 PT peppermint oil glands for enhancing expression of plant oil gland cell
 PT proteins -
 XX
 XX Claim 1; Page 165; 251pp; English.
 CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
 CC correspond to all or part of a mRNA molecule expressed in plant oil
 CC gland cells, especially peppermint and plant oil glands that produce
 CC terpenoid essential oils and resins. The nucleic acids are useful for
 CC genetically mapping a plant genome for genes expressed in plant oil
 CC gland cells and to suppress (for example by antisense suppression) or
 CC enhance their expression (for example by genetically transforming a
 CC plant cell with a replicable expression vector that expresses one or more
 CC proteins naturally expressed in plant oil gland cells). The nucleic acids
 CC are also useful for recombinant expression of plant oil gland proteins
 CC required for terpenoid essential oil and/or resin production in bacterial

CC and/or yeast cells.
 XX SQ Sequence 459 BP; 106 A; 141 C; 119 G; 93 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,53e-15 Length: 459
 Score: 377.00 Matches: 73
 Percent Similarity: 78.73% Conservative: 14
 Best Local Similarity: 61.86% Mismatches: 23
 Query Match: 21.36% Indels: 8
 DB: 22 Gaps: 1
 US-09-995-938a-7 (1-336) x AAH87878 (1-459)
 QY 23 ArgLYsProSerTYrArgGluArgGluAenAenArgArgGluArgArgArgAla 42
 DB 130 AGCTGCTCCGACATGAGAGAGCGGGGAAACAAGCGAGGAGCGCGCCCGCGCG 189
 QY 43 ValAlaAlaValIleTYrThrGlyLeuArgAlaGlnGlyAspTYrAenLeuProLYsHis 62
 DB 190 ATCGCCGCGAAGATCTTCTCGGCTGAGAGATGTACGGCACTTCAAGCTCCCAAGCAC 249
 QY 63 CySAspAenAenGluValLeuLYsAlaLeuCySValGluAlaGlyTYrValValGlu 82
 DB 250 TCGCACACAAAGAAAGTGTCTCAAGGCTCTCTGATGAAGCTGGCTGCTCATCAAGA 309
 QY 83 AspGlyThrThrTYrArgLYsGlyCyAlpProLeuProGlyGluLeaGlyThrSer 102
 DB 310 GACGGCACCACTTACAGAAAGGATGCAAGCTGTG----- 345
 QY 103 SerArgValThrProTYrSerSerGluAenGlnSerProLeuSerSeraIlePheGlnSer 122
 DB 346 GAACGTATGATATCATAGCTTCGGCAACAGTCAGTCTTGACCGCATATCAACCAAGC 405
 QY 123 ProIleProSerTYrGlnValSerProSerSerSerSerPheProSerProSer 140
 DB 406 CCTGAGAGCTCTTTTAAACCAAGTCTGATCTTCTTTTCTTCTGCTAGCCCTTC 459
 RESULT 9
 ABX22943
 ID ABX22943 standard; cDNA; 413 BP.
 XX
 AC ABX22943;
 XX
 DT 10-FEB-2003 (first entry)
 DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #5000.
 XX
 KW Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
 KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
 KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
 KW complex carbohydrate; gene replacement therapy; immunosuppressive;
 KW anti-inflammatory; antiarthritic; antibacterial; cerebroprotective;
 XX antiasthmatic; vasotropic.
 OS
 XX Homo sapiens.
 XX
 PN US2002110548-A1.
 XX
 PD 15-AUG-2002.
 XX
 XX 11-JUN-2001; 2001US-0878574.
 PF
 XX 22-NOV-1996; 96US-0753233.
 PR 03-DEC-1997; 97US-0984246.
 PR 09-SEP-1998; 98US-0148674.
 PR 14-JUN-1999; 99US-0333177.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Sullivan F, Kriz R, Kumar R;
 XX
 DR WPI: 2003-066673/06.


```

FT      misc_feature /note= "Mutational hot spot"
FT      907...908 /tag= Q
FT      /note= "Mutational hot spot"
FT      misc_feature 954 /tag= I
FT      /note= "Mutational hot spot"
XX      WO200177380-A2.
XX      PD
XX      18-OCT-2001.
XX      PF
XX      10-APR-2001; 2001WO-GB01622.
XX      PR
XX      10-APR-2000; 2000GB-0008801.
XX      PA
XX      (MEDI-) MEDICAL RES COUNCIL.
XX      PI
XX      Wright A;
XX      WPI; 2001-663057/76.
XX      DR
XX      PT
XX      Diagnosing disease or predisposition to disease, associated with
XX      disease causing mutations in retinitis pigmentosa GTPase regulator gene
XX      by genotyping ORF15 of the gene, and determining presence of mutations
XX      -
XX      PS
XX      Claim 6; Page 90; 10pp; English.
XX      CC
XX      The present invention relates to a method for diagnosing disease or
XX      CC      predisposition to a disease, associated with a disease causing
XX      CC      mutations in a retinitis pigmentosa GTPase regulator (RPGR) gene
XX      CC      involves genotyping a RPGR gene, and determining whether the genotype
XX      CC      comprises a disease causing mutation(s), where the risk genotype is
XX      CC      present within open reading frame (ORF)15 of the RPGR gene. The method
XX      CC      is useful for detecting a certain disease state e.g., X-linked
XX      CC      retinitis pigmentosa (XLRP). The kit is useful for detecting and
XX      CC      measuring disease causing mutations in biological fluids and tissues
XX      CC      and for localizing mutation in tissue. The mutant RPGR gene is useful
XX      CC      in gene therapy techniques and for screening agents capable of
XX      CC      affecting the expression of the sequences and/or the biological
XX      CC      activity of mutant RPGR. They are preferably useful for identifying
XX      CC      agonists and antagonists of RPGR. The mutant RPGR gene is also useful
XX      CC      in identification of potential pharmaceutical targets in high
XX      CC      throughput screening assays and forensic analysis. The present sequence
XX      CC      is the mutational hot spot sequence of human RPGR exon ORF15 DNA.
XX      CC      XX
XX      SQ      Sequence 1050 BP; 428 A; 15 C; 595 G; 12 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 0.0345 Length: 1050
XX      Score: 171.50 Matches: 69
XX      Percent Similarity: 44.09% Conservative: 28
XX      Best Local Similarity: 31.36% Mismatches: 77
XX      Query Match: 9.72% Indels: 47
XX      DB: 22 Gaps: 10
XX
US-09-995-938A-7 (1-336) x AAD21685 (1-1050)
Oy      93 ProLeuProGluGluIleAlaGlyThrSerSerArgValThrProTyr---SerSerGln 111H
Db      735 CCTCTTCCTCCCTCCCATTTCTCT-TCTCTCTTTTCCCTCCTCCCTTCATCTCCCT 677H
Oy      112 AsnGlnSerProLeuSerSerAlaIadheGlnSerProIleProSerTygIlnValSerPro 133H
Db      676 TCCTCTTCTCTCT--TCTCTCTCTCCCTCCCTCCCTCTTCCCTTCC-----TCTCCT 622H
Oy      132 SerSerSerSerPheProSerProSerArgGluProAsnAsnAmMetSerSerThr 151H
Db      628 TCCTCTCTCCCTTCTCTCTCTCTCTCTCTCTCTTCTCTCTCCCTCCCTCTCTCTCTCTCT 565H
Oy      152 PhePheProPheLeuArgAsnGlyGlyIleProSerSerLeuProSerLeuArgIleSer 171H

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 20, 2003, 01:10:37; Search time 2483 Seconds

(without alignments)
3288.886 Million cell updates/sec

Title: US-09-995-938A-7
Perfect score: 1765
Sequence: 1 MTSGATSTSTAAAAA.....IHDVGMEDLTLNGKARG 336

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US0995998/runat.19122003.142606.4249/app.query.fasta.1.519
-DB=EST -OPWT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFC=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0995998 @CGN 1.1 2810 @runat.19122003.142606.4249 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUEYRY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estsba:*
2: em_estchum:*
3: em_estcin:*
4: em_estcmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estcro:*
8: em_hlc:*
9: gb_estcl:*
10: gb_estc2:*
11: gb_hlc:*
12: gb_estc3:*
13: gb_estc4:*
14: gb_estc5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rpd:*
26: em_gss_phg:*
27: em_gss_vrt1:*
28: gb_gssb1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	913.5	51.8	772	13	CB292898
2	901	51.0	796	14	CB292898 UCRC01.0
3	885.5	50.2	789	10	BU238047
4	872.5	49.4	765	10	BG127474
5	871	49.3	768	12	BG126322
6	857	48.6	585	12	BG887699
7	838.5	47.5	762	12	AV551486
8	825.5	46.8	766	10	BG593827
9	778	44.1	658	10	BG130895
10	777.5	44.1	643	14	CA934625
11	776.5	44.0	619	9	AV440980
12	763.5	43.3	941	28	B12579
13	759.5	43.0	648	9	AV545003
14	752	42.6	602	29	B2476824
15	746	42.3	639	9	AV822043
16	743	42.1	1001	12	BG840094
17	732.5	41.5	864	14	CB893187
18	732	41.5	744	12	B1921662
19	730.5	41.4	773	28	BH673471
20	729	41.3	685	9	AV782657
21	720.5	40.8	840	14	CB892700
22	719.5	40.8	729	13	BU027742
23	716.5	40.6	717	13	BU027177
24	711.5	40.3	708	10	BF480563
25	709.5	40.2	724	13	BU024180
26	696.5	39.5	717	13	BU025606
27	692	39.2	553	10	BE450956
28	689	39.0	658	13	BQ138490
29	687	38.9	801	12	B1433025
30	684	38.8	590	10	BE343539
31	680.5	38.6	766	29	B2476333
32	673.5	38.2	610	10	BE803614
33	671	38.0	817	14	CB893419
34	660	37.4	614	9	AW317351
35	655.5	37.1	677	9	AW299184
36	654	37.1	625	10	BE023315
37	649	36.8	511	10	BF273350
38	645.5	36.6	629	13	BQ407744
39	644	36.5	579	12	BM062278
40	635.5	36.0	590	9	AV822219
41	626	35.5	511	28	BH597004
42	626	35.5	539	9	AV442375
43	623.5	35.3	679	9	AW689777
44	622	35.2	676	13	BQ917002
45	620	35.1	580	14	CD481007

ALIGNMENTS

RESULT 1
CB292898
LOCUS
DEFINITION
albedo cDNA library Citrus sinensis
ACCESSION
CB292898
VERSION
CB292898.1
KEYWORDS
SOURCE
ORGANISM
Citrus sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

FEATURES
source

Location/Qualifiers
1..796
/organism="Deecuratinia sophia"
/mol_type="mRNA"
/db_xref="taxon:89411"
/clone="De01_14f06"
/tissue_type="leaf, stem"
/dev_stage="1-month seedlings - 1 cm tall - 8 leaf"
/clone_1lb="De01_AARC_ECORC_cold_stressed_flixweed_seedling"

/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: Eco RI; Site 2: Xho I; Plants were grown for 1 month at 200C/16 hrs light/day (average 8 leaves, 1 cm tall, weight 0.02g/plant). Then they were exposed to 20C, 12 hrs light/day, for 1 week. Library prepared by C. Piche using Stratagene Kit."

BASE COUNT 221 a 150 c 211 g 207 t 7 others
ORIGIN

Alignment Scores:

Pred. No.: 9.08e-58 Length: 796
Score: 901.00 Matches: 178
Percent Similarity: 87.56% Conserved: 5
Best Local Similarity: 85.17% Mismatches: 22
Query Match: 51.05% Indels: 4
DB: 13 Gaps: 3

US-09-995-938a-7 (1-336) x BU238047 (1-796)

QY 129 ValSerProSerSerSerSerPheProSerProSerArg---GlyGluProAnaAnaAn 147
DB 796 GTCAGTCTCATGCTTCATCATCTCCAGTCTTCGAGTGGCAGCCACACAC-- 740
QY 148 MetSerSerThrphePheProPheLeuArgAnGlyGlyLeuProSerSerLeuProSer 167
DB 739 ---ATCTCCACAAATCTCCGCTTCCTCAGAAATGGCGAATCCCTTCATGCTTCCTCG 683
QY 168 LeuArgGleSerAnSerCyProValThrProValSerSerProThrSerLeuAn 187
DB 682 CTTAGGATCTCAACACAGTGTCCAGTCACTCCAGCTCATCTCAAAATTTCAAAAC 623
QY 188 ProLeuProLeuProAnaThrGluSerLeuLeuArgGlnSerMetAlaLeu---AlaLys 206
DB 622 CCAAAACCAATGGCCACTGGAGTCTTTTACCAACCAATCCAGCCATGGCTGCTAAA 563
QY 207 GlnSerMetAlaSerPheAnSerProPheThyAlaValSerAlaProAlaSerProThr 226
DB 562 CAATCGATGACATTTTAACTATCCATTTTAAACCCGATTCGACCTGCCATCCAC 503
QY 227 HisArgHisGlnPheHisThrLeuAlaThrLeuProGluCyAspGluSerAspSerSer 246
DB 502 CATCATGCGCAGTTCATGCCCCCGCTACTATCTGTAATGTGATGAGTCTGACTCTCC 443
QY 247 ThrValAspSerGlyHisTrpPheSerPheGlnLysPheAlaGlnGlnGlnProPheSer 266
DB 442 ACTGTGACTCGGCTCATGTGATAAGCTTTCAAAAGTTCTCAACAGCCATTCATGTC 383
QY 267 AlaserMetValProThrSerProThrPheAnSerValLysProAlaProGlnGlnMet 286
DB 382 GCGCTCTGCTGCCACCTCTCCAACTCATCTTAAGTAAGAACCCGCTCCACCAATTA 323
QY 287 SerProAnaThraAlaPheGlnGluLeuGlnSerSerGluPheLysPheGluAn 306
DB 322 TCTCCAAACACTGCTGCAACCAAGAGATGTGTCAAACTCGGAATTTAAATTCAGAAC 263
QY 307 SerGlnValLysProTrpGluGlyGluArgGlnHisAspValGlyMetGluAspLeuGlu 326
DB 262 AGCCAAATTAAGCATGGAGAGAGAGATCATGATGCTATGAGATGAGATCTAGAG 203
QY 327 LeuThrLeuGlyAnGlyLysAlaArg 335
DB 202 CTCACACTTGGAAACGGCAAGCTCTGT 176

RESULT 3

LOCUS BG127474

DEFINITION BG127474 789 bp mRNA linear EST 31-JAN-2001

ACCESSION EST473120 tomato shoot/meristem Lycopersicon esculentum cDNA clone

VERSION CIOF16J21 5' sequence, mRNA sequence.

KEYWORDS BG127474.1 GI:12627662

SOURCE EST.

ORGANISM Lycopersicon esculentum (tomato)

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 789)
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,
Hansen, C., Romling, C. and Tanksley, S.

Generation of ESTs from tomato shoot/meristem tissue

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.

Location/Qualifiers

FEATURES
source

1..789
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CIOF16J21"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/clone_1lb="tomato shoot/meristem"
/note="Vector: Bluescript SK(-); Site 1: EcoRI; Site 2: XhoI. Small expanding leaves from the growing tip were taken from greenhouse plants (4-6wks old TA496). Tissue was immediately frozen in liquid nitrogen."

BASE COUNT 218 a 186 c 160 g 225 t
ORIGIN

Alignment Scores:

Pred. No.: 1.3e-56 Length: 789
Score: 885.50 Matches: 181
Percent Similarity: 76.62% Conserved: 32
Best Local Similarity: 65.11% Mismatches: 46
Query Match: 50.17% Indels: 19
DB: 10 Gaps: 7

US-09-995-938a-7 (1-336) x BG127474 (1-789)

QY 33 AsnArgArgArgGluArgArgArgAlaValAlaLysIleTyThrGlyLeuArg 52
DB 3 AATAGAGAGAGGAAAGAGAGAGAGCAATACAGTAAAGATTATAGATGATTAAGA 62
QY 53 AlaGlnGlyAspTyraLeuLeuProLysHisCysLeuPheAnGluValLeuLysAlaLeu 72
DB 63 GCACAGGGGAATATATATCTTCTTAAACATTTGATTAACAATGAGGTTTGAAGCTCTT 122
QY 73 CysValGluAlaGlyTrpValValGluGluAspGlyThrTrpTyraLysGlyCysLys 92
DB 123 TGTGTTAACTGGATGATGATGTTGACCTGATGAATCTTATATAGAAAGGAGTCAAG 182
QY 93 ProLeuProGluGluIleAlaGlyThrsSerSerArgValThrProTyrsSerGlnAn 112
DB 183 CCAACTCAATGAGATTTGAGGACCTTCAGCCAAATTAAGCCCAAGTCTTCAGAAAT 242
QY 113 GlnSerProLeuSerSerAlaPheGlnSerProIleProSerTyraGlnValSerProSer 132
DB 243 CCAAGTCTCTCCCTTCTTACTTGTAGCCCGAATTCATCTTCAAGATTAGTCAACA 302
QY 133 SerSerSerPheProSerProSerArgGlyGluProAnaAnaMetSerSerThrphe 152

Db 303 TCCTGCTTTTCCCAAGTCATCTGCTGGTATGCT-----ACATGTCGTCACATCA 356

Qy 153 PheProPheLeuAAsnGlyGlyLeuProSerSerLeuProSerLeuAArgLeuSerN 172
|||
Db 357 TTTCGATTTCATC---CATAGTTCATCTCCCTTGCTGCATCCACCATTGCAATATCAAC 413

Qy 173 SerCySProValThProProValSerSerProThSerLyAsnProLyProLeuPro 192
|||
Db 414 AGTGCCTCTGTACACCACTCTTTCATCAACCACT---AGAGTCCCTTAAGCAGATTTT 470

Qy 193 AsnTrpGluSerIleAlaIleGlnSerMetAlaIleAlaIleGlnSerMetAlaSerPhe 212
|||
Db 471 AATCTTGAGACT-----TTGGCTAGAGACTCATGTCTGCTCTA 509

Qy 213 AsnTyTrProPheTyAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHis 232
|||
Db 510 AATATCCCTTCTTCTTGCTGCTTGCACCCCACTACCTCCCACTCGAGTCAGCATTC--- 566

Qy 213 ThrLeuAlaThrIleProGluCysAspGluSerAspSerSerThValAspSerGlyHis 252
|||
Db 567 ACTCTGCTACAAATACCAAGAGTACAGATGTGATTCATCCACCATGATTCCTGGCAG 626

Qy 253 TrpIleSerPheGlnIlePheAlaGlnGlnGlnPheSerAlaSerMetValProThr 272
|||
Db 627 TGGATGAGACTTTCAAAAGATGCA-----GCCAAATGGATCTCTACT 668

Qy 273 SerProThrPheAsnLeuValIleProAlaProGlnGlnMetSerProAsnThrIleAla 292
|||
Db 669 TCTCCGACTTTTAATCTTATTAAAGCTGTAGCTCAGAGAAATCTTCTTAATGATATGATC 728

Qy 293 PheGlnGluIleGlyGlnSerSerGluPheLyPheGlnIleAsnSerGlnValIle 310
|||
Db 729 ATCGACAAG--GCTAAGAGCATTTGAATTTGACTTTGAGAAATGATATCAGTTAAG 779

RESULT 4
Bg126322

LOCUS Bg126322 765 bp mRNA linear EST 31-JAN-2001

DEFINITION EST171568 tomato shoot/meristem Lycopersicon esculentum cDNA clone

CTOP12A3 5' sequence. mRNA sequence.

ACCESSION Bg126322

VERSION Bg126322.1 GI:12626510

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

EXTRACTS Eukaryote; Vitisidplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 765)

REFERENCE van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Uterback,T.,
Hansen,C., Roning,C. and Tankley,S.
Generation of ESTs from tomato shoot/meristem tissue

TITLE Unpublished

JOURNAL Contact: CGCI

COMMENT Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
Location/Qualifiers

FEATURES
Source 1..765

ORIGIN

[illegible]

US-09-995-938A-7 (1-336) x BG126322 (1-765)

[illegible]

ACCESSION	BG887699
VERSION	BG887699.1
KEYWORDS	GI:14264785
SOURCE	EST.
ORGANISM	Solanum tuberosum (potato)
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.
AUTHORS	van der Hoeven,R., Bezzerides,J., Ewing,E., Cho,J., Chiemingo,A., Bougri,O., Buell,C.R., Roming,C., Tankley,S. and Baker,B.
TITLE	Generations of ESTs from dormant potato tubers
JOURNAL	Unpublished
COMMENT	Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-array@tigr.org This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL: http://genome.arizona.edu/orders/ Seq primer: M13P-R. Location/Qualifiers
FEATURES	
source	1..768 /organism="Solanum tuberosum" /mol_type="mRNA" /cultivar="kennebec" /db_xref="taxon:4113" /clone="cstD6U10" /tissue_type="dormant tuber" /dev_stage="one month post-harvest" /lab_host="SOLR" /clone_idb="cstrD" /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; This library targets genes expressed in dormant tubers. This library was made from sections of dormant tuber, avoiding the buds and epidermis. Tubers were stored for one month post-harvest at 4°C. The tuber was peeled, well away from the surface. Then it was chopped into 1-2 mm cubes and immediately frozen in liquid nitrogen. This library is noted as P4 in Tanksley lab notebooks."
BASE COUNT	220 a 209 c 133 g 206 t
ORIGIN	
Alignment Scores:	
Pidm. No.:	1,53e-55 Length: 768
Score:	871.00 Matches: 183
Percent Similarity:	77.07% Conservative: 22
Best Local Similarity:	68.80% Mismatches: 41
Query Match:	49.35% Indels: 21
DB:	12 Gaps: 6
US-09-995-938A-7 (1-336) x BG887699 (1-768)	
Oy	25 ProSeRTPaRGluARgluaBnaandrgATgAGTguARGArgrAlaVala1a 44
Db	2 CCATCATGAGAGAAGAAAGAGATAAACGAGAGAAGAGAAGAGAGACTGTACT 61
Oy	45 AlAlYbIlEtTYThChLYleuARglagInGLyaBPYYrAnleuProLyehIeCyAsp 64
Db	62 GCTAAGATTAAACGGTTTAAGACTCAAGAACTTAACACTTCCTAAACAAGTGTA 121
Oy	65 AenaENGluValleuLeuAlaleucCyVaIGlualAGLTTPVaIVaIGluGlunaspGL 84
Db	122 AACATGAAGTCTTAAAGCTCTTGTACTGAAGCTGTTGATGATCGTTCAACCTAAGT 181
Oy	85 ThTrHrTYrARGLyGLYCAlalysProlEuPrOgILjulaGLYThIsErSerArg 104
Db	182 ACCACTTATCCGAAGGATGCAAGCAACCAGATGAGATGGAGGCACCTCAACAAC 241
Oy	105 ValThrProTyLserSerClnhngInserProlEserSerAlaphedInsErProle 124
Db	242 ATCAGCCCAAGTTCTTCACGGATCCCAAGTCCCGCTATCACTACTTGTGTCGCCAAT 301

Oy		125	ProserYrGlnAlaSerProSerSerSerSerPheProSerProSerAAGGlyGluPro	144
Dd		302	CCATCTTATACGCCAACTCCAACTCCTCTTTTCCCCAGTCATCCGTGCGTAGCT	361
Oy		145	AAsnAAsnMeSerSerThrPhePheProPheLeuArgAnGlyGlyIleProSerSer	164
Dd		362	-----AACATGCATCACATCCATATCTTTCTTCATAT--GTCATCCCTCATCC	412
Oy		165	LeuProSerLeuArgIleSerAAsnSerCyseProValThrProProValSerSerProThr	184
Dd		413	CTTCTCTCATTTACGAATATCAACAGAGCCCTGGTAATCCACTCTTCATCCACAAT	472
Oy		185	SerIysAnnProlYsrProLeuProAnProAnTrpGluSerIleAlaIysGlnSerMetAlaIle	204
Dd		473	---AGGATCCTTAAGCAACATTCAATTTAGAAGCT-----TTG	508
Oy		205	AlaIysGlnSerMetAlaSerPheAenTyProPheTyAlaValSerAlaProAlaSer	224
Dd		509	GCCAAAGATCATGTTTGCTTTAAACATCCCTTTCTTCTGCTTCAGCCCAAGCAGC	568
Oy		225	ProThrIlaArgHisGlnPheHicThrLeuAlaThrIleProGluCyAspGluSerAs	244
Dd		569	CCAACTAGAGTCAGCGCTTT---ACTGCTCCCAACTATACCAGAGTGATGATCTGA	625
Oy		244	pSeSerThrValAspSerGlyHisTrpIleSerPheGlnLysPheAlaGlnGlnInPr	264
Dd		626	CTCATCTTACATGATTCAGCGCAGTGATCACTTTCAAAAGATC-----	671
Oy		264	oPhSeValAséMerValProThrSerProThrPheAsnLeuValLyseProAlaProGl	284
Dd		672	-----GCCGGCAGCTTCCACTTCCACA-TTTAATCTGTAAACCTGCGCTCA	723
Oy		284	nGlnMetSerProAn	289
Dd		724	GCCACTTCGTCGTAAT	739
RESULT 6				
AVS51486				
LOCUS		AVS51486	585 bp	mRNA linear EST 06-SEP-2000
DEFINITION		AVS51486 Arabidopsis thaliana roots Columbia Arabidopsis thaliana		
ACCESSION		AVS51486		
VERSION		AVS51486.1	GI:8722899	
KEYWORDS		EST.		
SOURCE		Arabidopsis thaliana (chale crees)		
ORGANISM		Arabidopsis thaliana		
REFERENCE		Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS		Spermatophytes; Magnoliophyta; eudicotyledons; core eudictor; rosids		
TITLE		; eurosids II; Brassicales; Brassicaceae; Arabidops.		
JOURNAL		1. (bases 1 to 585)		
MEDLINE		Asamizu,E.; Nakamura,Y., Sato,S. and Tabata,S.		
PUBMED		A large scale analysis of cDNA in Arabidopsis thaliana: Generation		
COMMENT		of 12,028 non-redundant expressed sequence tags from normalized and		
		size-selected cDNA libraries		
		DNA Res. 7, 175-180 (2000)		
		Contact: Erika Asamizu		
		The First Laboratory for Plant Gene Research		
		Kazusa DNA Research Institute		
		Yana 153-3, Kisarazu, Chiba 292-0812, Japan		
		Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.		
FEATURES		Location/Qualifiers		
Source		1..585		
		/organism="Arabidopsis thaliana"		
		/mol_type="mRNA"		
		/strain="Columbia"		
		/db_xref="taxon:3702"		
		/clone="R2127d10R"		
		/tissue_type="roots"		
		/clone_tib="Arabidopsis thaliana roots Columbia"		
		/note="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2:		
		XhoI"		


```
Db          CTTGCTACAAATCCAGGTGCATGTCATTGAATCATCACCATTGATTTGGCCAGTGG      625
Oy         254   ILeSerPhegInLysPheAlaGLnglnGnPropHeSeRaLaSeR-MeValProThrSe    273
                ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db        626 ATGAGCTTTCAAAAGTAGTA-----GCCAATGGGATCGCTACTTC       667

Oy         273 rProThrHemLenUeuVallLySProlALPrOGlInImetSrProAdenThraLAalAh     293
                |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db        668 TCCGACITTTATCTTATTATAAGCCCCTAGTGCTACAAGAATTCCTTCTAATGATATGATCAT      727

Oy         293 eGlNgLuIleGIynSereSGluPhalySpheglu    305
                :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db        728 CGACAg--GATAAGAcATTGaGcTTGAttttgaG      761


RESULT 8
Bg593827                               766 bp           mRNA             linear EST 07-MAR-2003
LOCUS Bg593827                                Solanum tuberosum cDNA clone CSTSSH7 S' sequence,
DEFINITION mRNA sequence.
ACCESSION Bg593827
VERSION Bg593827.1 GI:13611967
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; Eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1..to 766)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,Y., Chimingo,A.,
Bongri,O.O., Buell,C.R., Romling,C.J., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potatoc-array@igrr.org
This clone can be obtained from the University of Arizona Genomics
institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13-F-R.
```

D	B		GAAAGAGAAGATAC--AGGAGAAGAGAGAAGAAGCAGAGACTGTAGTGGCTGAAGATTAT	61
OY		49	ThrlglyLeuAlrGalaglInglyAseTryAenLeuProLyhiScySaapAanBengluVal	68
D	B	62	ACTGGTTTAAGAGCTCAAGGAATACTAATCTTCCTAAACACTGTGATACAAAGAAGTT	121
OY		69	LeuIysAlaLeuCySeValGluaAgLTyrValValGlugluAepSgLYthrThrTYArX	88
D	B	122	CTTAAAGCTCTTGTTACTGAAGCTGGTGGATGCCTTAACCTGATGGTACCATTATCGC	181
OY		89	LysglICyShsPProleuProglYglulIealaglyThrsEserAargValThProTyx	108
D	B	182	AAGGAGATGCCAACCCCGATGAGATGAGAGCATTCACAAACAATCACGCCCAAGT	241
OY		109	SerSerGlnaBngInSerProLeuSerSerAlaPhelGlnSerProIleProSerTYGln	128
D	B	242	TCTTCACGGCAGTCCAATCCCCCGCTCATCTACTTCTGAGCCCAATTCATCTTATCAG	301
OY		129	ValSerProSenSerSerSerPhePheSerPheSerPheSerLrglgLuPProAnaAnaMet	148
D	B	302	CCAAGTCCAACTTCCTCTTCTTCCCAGGCCAATCCGTCGTATGCT----AACATG	355
OY		149	SerSerThrPhePhePheProPheLeuArgAsnGLylIleProSerSerLeuProSerLeu	168
D	B	356	TCATTCACATCCATATCTTCTTCTCCATAT--GTCAATCCCTTATCTCTTCATTA	412
OY		169	ArgIleSerAnSerCySProValThlrProProValSerSerProThrSerLysAnBro	188
D	B	413	CGAATATCAAACGATGGCCCTGTAACTGCACCTCTTCATCACCAACT--AGCATCTCT	469
OY		189	LysProLeuPProaantTPglUserlIeAlaLysGlnSerMetAlaIleAlLysGlnSer	208
D	B	470	AAGCAAACTTCAAATTTAGAGACT-----TTGGCCMAAGAATCA	508
OY		209	MetaIasErPheaenyYrProPheTYRAlaValSerAlaProAlaSerProThrhiaX	228
D	B	509	ATGTTTGCTTTAAACATCCCTTCTTCTTGCTGCTTCAGCCCAAGAACCCAACTAGGGGT	568
OY		229	HieGlnPheHistrLeuAlaThrIleProgluCysApBgIuseAaspSerSerThryAl	248
D	B	569	CAGGTTTT-----	577
OY		249	AspSerGlyHistrPlieserPheGlnLysPheAlaGlnGlnInProPheSerAlaser	268
D	B	577	-----	577
OY		269	MetValProthnsErProThrPheAsnLeuValLysProAlaProGlnGlnMetSerPro	288
D	B	578	-----ACTCCTCCAACTTTAATCTGTAAACCTGTGCTTCAGCCACTTGTCTCT	628
OY		289	AanthraIalaIaphelGlnluIleGlyGlnSerSerGluPheLysPheGluAnSerGln	308
D	B	629	AATGAT---ATGATCACAGACAAAGGTAAAGCATAGACTTCGACTTGAATAATGTATCA	685
OY		309	ValIysProTPglUglYgluArgILehIaSpValGIymetGluAepLeuGluLeuThr	328
D	B	686	GTCAAAGCATGGGAAGGGGAAGGATTCACGAATGTAGATTGATGTGATGTGGACTCAC	745
OY		329	LeuGlyAenGlyLysAlaArg	335
D	B	746	CTTGAAAGTGGCAATGTCTGC	766
 RESULT 9 BGJ30895 658 bp mRNA linear EST J1-JAN-2001				
DEFINITION ES1463787 tomato crown gall Lycopersicon esculentum cDNA clone				
ACCESSION BGJ30895				
VERSION BGJ30895.1 GI:12631083				
KEYWORDS EST.				
SOURCE Lycopersicon esculentum (tomato)				
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				


```

Db      122 GGGGAT-----AACATGCGCTTTTATCTCTTCATCTCTTCAGT---GCCATT 172
Qy      162 ProSerSerLeuProSerLeuArgIleSerAsnSerCybProValThrProProValSer 181
Db      173 CCATGCTCTTCTCTCTCTCTCTTCGATCTCAACAGTCGCGCTGTACACCCACCTCTCG 232
Qy      182 SerProThSerLeuAsnProLysProLeuProAsnThrGluSerIleAlaLysGlnSer 201
Db      233 TCCCGGACCTCAAGAAACCCCAACCAATTCACAGTGGAC-----274
Qy      202 MetAlaIleAlaLysGlnSerMetAlaSerPheAsnTyrrProPheTyrrAlaValSerAla 221
Db      275 ---TTTATTGCCAACAATCATCGCTCTCTTATTGTTACCATTTAAGCGGTCTCGCT 331
Qy      222 ProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrIleProGluCyAsp 241
Db      332 CCGGCTGACCCCACTCACCGT---CAGTTTCATCTCTCAGACCATTAATACGATGAT 388
Qy      242 GluSerAspSerLeuThrValAspSerGluHisThrIleSerPheGlnLysPheAlaGln 261
Db      389 GAGCTGATCATCTCACTGTTGATGTCGTGATGATGATGATGATGATGATGATGATGATGAT 445
Qy      262 GlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsnLeuValPro 281
Db      446 -----CTTCTGTGCTGCGACGATGCCAACCCTCTCTACTTAACTTTGTGATACC 499
Qy      282 AlaProGlnIleMetSerProAsnThrAlaAlaPheGlnGlnIleGlnSerSerGlu 301
Db      500 GTGGCTCAGCAATTTGCTCAACAATTTGCTCAAGAGATGA---GTGCCCATGAT 556
Qy      302 PheLysPheGlnLysSerGlnValLysProThrGluGluArgIleHisAspValGly 321
Db      557 TTGGAGTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 616
Qy      322 MetGluAspLeuGluLeuThrLeuGly 330
Db      617 TTAGATGATCTAGAGCTCACCTTGA 643

RESULT 11
AV440980      619 bp      mRNA      linear      EST 14-NOV-2000
LOCUS      AV440980 Arabidopsis thaliana above-ground organ two to six-week
DEFINITION      Old Arabidopsis thaliana cDNA clone Ap218b09_f3', mRNA sequence.
ACCESSION      AV440980
VERSION      AV440980.1 GI:7611357
KEYWORDS      EST.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 619)
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093
10907847
COMMENT      Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yena 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1. 619
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="Ap218b09_f"
/label_type="aboveground organs"
/dev_stage="two to six-week old"

```

```

BASE COUNT      176 a      129 c      148 g      166 t
ORIGIN
Alignment Scores:
Pred. No.:      1,41e-48      Length:      619
Score:      776.50      Matches:      147
Percent Similarity:      92.07%      Conservative:      4
Best Local Similarity:      89.63%      Mismatches:      12
Query Match:      43.99%      Indels:      1
DB:      9      Gaps:      1

US-09-995-938a-7 (1-336) x AV440980 (1-619)
Qy      172 AsnSerCybProValThrProProValSerSerProThrSerLysAsnProLysProLeu 191
Db      618 AACAGTCTCTCTGCTCACTCCACGATGTCATCCCACTTTAGAAACCCCAACCATTTG 559
Qy      192 ProAsnTrpGluSerIleAlaLysGlnSerMetAlaIle--AlaLysGlnSerMetAla 210
Db      558 CCACTTGGAGATCTTTTACCAACAATTCATGTCATGCGTGTAAACAGTCATGACT 499
Qy      211 SerPheAsnTyrrProPheTyrrAlaValSerAlaProAlaSerProThrHisArgHisGln 230
Db      498 TCTTGAACCTAACCCGTTTATGCGGTGTGCACTGCACTGCACTCATCATCATGCCAG 439
Qy      231 PheHisThrLeuAlaThrIleProGluLysAspGluSerAspSerSerThrValAspSer 250
Db      438 TTCATCTCTCCGCTACTTAACTGAAATGATAGTGTGACTCTTCCACTGTGATCT 379
Qy      251 GHisThrIleSerPheGlnLysPheAlaGlnGlnIleProPheSerAlaSerMetVal 270
Db      378 GGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 319
Qy      271 ProThrSerProThrPheAsnLeuValLysProAlaProGlnIleMetSerProAsnThr 290
Db      318 CCAACTCGCTACCTTCAATCTCGTGAACCTGCACACGACCAATGCTTCCCAACACA 259
Qy      291 AlaAlaPheGlnGlnIleGlnSerSerGluLysPheGlnLysAsnSerGlnValLys 310
Db      258 GCAGCAATCAAGAGATGTCGAAGCTCCAGTTTAAAGTTGAGAACGACCAAGTTAAG 199
Qy      311 ProThrGluGluArgIleHisAspValGlyMetGluAspLeuGluLeuThrLeuGly 330
Db      198 CCATGGAAAGGAGAGATCCATGATGCTATGAGAGATCTTAGACTCAGCTTGA 139
Qy      331 AsnGlyLysAla 334
Db      138 AATGTAAGCT 127

RESULT 12
B12579      941 bp      DNA      linear      GSS 14-MAY-1997
LOCUS      F22H5-T7 IGF Arabidopsis thaliana genomic clone F22H5, genomic
DEFINITION      survey sequence.
ACCESSION      B12579
VERSION      B12579.1 GI:2093699
KEYWORDS      GSS.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 941)
Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shim, P., Sun, H. and
Ecker, J.
BAC End Sequences at ATGC
TITLE      Unpublished: F22H5-SP6
JOURNAL      Other: GSS: F22H5-SP6
COMMENT      Contact: Ecker J.

```

Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@genome.bio.upenn.edu

Seq primer: 77
Class: BAC ends
High quality sequence start: 94
High quality sequence stop: 469.

FEATURES

source

location/Qualifiers

1..941
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F22H5"
/bex="hermaphrodite"
/clone_1db="IGF"
/note="Vector: BeLoBACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"
BASE COUNT 230 a 213 c 157 g 324 t 17 others
ORIGIN

Alignment Scores:

Pred. No.: 2,198-47 Length: 941
Score: 763.50 Matches: 159
Percent Similarity: 81.04% Conservative: 12
Best Local Similarity: 75.36% Mismatches: 35
Query Match: 43.26% Indels: 5
DB: 28 Gaps: 1

US-09-995-938a-7 (1-336) x B12579 (1-941)

Qy 89 LyseGlyCyblyrProleuProGlyLutleaglYthnSerSerArgValThrProTyr 108
Db 312 CAGAGTATCAACCTTCTACCTGTGAGATAGCTGGAGACTCTCGAATTACTCCAT 371
Qy 109 SerSerGlnAnGlnSerProleuSerSerSerAlaPheGlnSerProIleProSerTyrGln 128
Db 372 TCATACAGAACGAGCCCTTCTTCAATCAGCTTCAAGTCCATCCATCCATCCAA 431
Qy 129 ValSerProSerSerSerSerPheProSerProSerArgGlyGluProAnAnAnMet 148
Db 432 GTTAGCCGCTCTTCTTCAATCCGGAATCTTCCGGGGGAAACAATAACAATG 491
Qy 149 SerSerThrPhePheProPheLeuArgAnGlyLysIleProSerSerLeuProSerLeu 168
Db 492 TCCTCTACATCTCTCTCTCTCTCAGAAATGCTGCATCTCTCTCTCTCTCTCTC 551
Qy 169 ArgIleSerAnSerCybProValThrProProValSerSerProThrSerLysAnPro 188
Db 552 AGAATCTCAACAGTGTTCAGATTACCAACCGGCTCTCAACGCACTTCAAGAAACG 611
Qy 189 LysProLeuProAnThrGlnSerIleAlaLysGlnSerMetAlaIleAlaLysGlnSer 208
Db 612 AANAGCTGCTCAATCGGAATCTATCGTAAGCAATCATTCGCAATCTCTAAACATTA 671
Qy 209 MetAlaSerPheAnThrProPheTyrAlaValSerAlaProAlaSerProThrHisArg 228
Db 672 ATGGCGTCTTTAAATATCTTCTATAGCGGTTCTGCNACTGCTAATCCACACATCTG 731
Qy 229 HisGlnPheHisThrLeuAlaThrIleProGlnCybAspGlnSerAspSerSerThrVal 248
Db 732 CAGCGTTTCATACCCCGCTACTAAACTGAATGATGAGCTGACCTTCCACTGT 791
Qy 249 AspSerGlyHisIleThrPhePheGlnLysPheAlaGlnGlnGlnProPheSerAl 267
Db 792 TGATTCGTGCTATGGGATTAATCTTCAAAAAGTTTGCAACANACACACATTCCTCG 851
Qy 267 aserMetVal---ProThrSerProThrPheAnLeuValLysProAlaProGlnGlnMe 286
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

Db 852 TCCTCTATGTGGCAACTCTCTCAATCTTCATCTGTGAAGTGGCCCCACATAT 911
Qy 286 tSerProAnThrAlaAlaPheGln 294
Db 912 GTCTCAATATATGGTGGCTTCAAG 936

RESULT 13

AV545003/c

LOCUS

DEFINITION

AV545003 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
cDNA clone R267c11F 3', mRNA sequence.

ACCESSION

AV545003

VERSION

AV545003.1

KEYWORDS

GI:8716417

SOURCE

Arabidopsis thaliana (chale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 648)

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.

A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries

DNA Res. 7, 175-180 (2000)

JOURNAL

MEDLINE

20363093

PUBMED

10907847

COMMENT

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yama 1532-3, Kiseazuru, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

location/Qualifiers

1..648

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="R267c11F"

/issue_type="roots"

/clone_1db="Arabidopsis thaliana roots Columbia"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 187 a 136 c 151 g 174 t

ORIGIN

Alignment Scores:

Pred. No.: 2,88-47 Length: 648
Score: 759.50 Matches: 144
Percent Similarity: 92.50% Conservative: 4
Best Local Similarity: 90.00% Mismatches: 11
Query Match: 43.03% Indels: 1
DB: 9 Gaps: 1

US-09-995-938a-7 (1-336) x AV545003 (1-648)

Qy 176 ValThrProProValSerSerProThrSerLysAnProLysProleuProAnThrGln 195
Db 648 GTCACTCCACAGTGTATCCCACTTCTAAGAACCCCAACCATTCCTACTTGGAA 589
Qy 196 SerIleAlaLysGlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPheAnThr 214
Db 588 TCTTTTACCAACATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
Qy 215 ProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeu 234
Db 528 CCGTTTATGCGGTCTGACCTCCACTGCTATCTACATCATCGCCAGTTCCATGCTCCG 469
Qy 235 AlaThrIleProGlnCybAspGlnSerAspSerSerThrValAspSerGlyHisIlePile 254
Db 468 GCTACTTATACCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 409
Qy 255 SerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerPro 274

Db 408 AGTTTCAAAAGTTTGACAAACAGCCATTCTGCTCTATGCGCACTGCGCT 349
 Qy 275 ThrPheAsnLeuVallyPheProAlaProGlnGlnMeSerProAnthraAlaPheGln 294
 Db 348 ACCCTCAATCTCTGAAACCTGCACCAACCAATTGCTCTCAACACAGACCAATCCAA 289
 Qy 295 GlnIleGlnGlnSerSerGlnPhePheGlnAsnSerGlnVallyPheProTyrGlnGly 314
 Db 288 GAAATGGTCAAAAGCTCCGAGTTTAAAGTTGAGAACAGCCAGTTAAAGCCATGGAGGG 229
 Qy 315 GlnArgIleHisAspValGlnMetGlnAspLeuGlnLeuThrLeuGlnIleVallyAla 334
 Db 228 GAGAGATCATCATATGCTGCTATGAGAGATCTAGAGCTACGCTTGAAATGGAAGCT 169
 RESULT 14
 B2476824/c 602 bp DNA linear GSS 13-DEC-2002
 LOCUS BONH01R.BQ_1.6_2_KB.tot Brassica oleracea genomic clone BONH01.
 DEFINITION genomic survey sequence.
 ACCESSION B2476824
 VERSION B2476824.1 GI:26779222
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eustosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 602)
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtowne@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..602
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BONH01"
 /note="Vector: pHOSt1; Site 1: BstXI; 1.6-2 kb sheared
 total DNA inserted into pHOSt1 using BstXI linkers"
 BASE COUNT 168 a 109 c 198 g 127 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,33e-47 Length: 602
 Score: 752.00 Matches: 158
 Percent Similarity: 82.69% Conservative: 14
 Best Local Similarity: 75.96% Mismatches: 22
 Query Match: 42.61% Indels: 14
 DB: 29 Gaps: 8
 US-09-995-938a-7 (1-336) x B2476824 (1-602)
 Qy 94 LeuProGlnGlnIleAlaGlnThSerSerArgValThrProTyrSerSerGln---Asn 112
 Db 600 CTGCTGTGATGACATGCTGATCTCTCTGAGCCACTCTTACTCTTCTTATAC 541
 Qy 113 GlnSerProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerProSer 132
 Db 540 CAAAGCCCT-----TTGAGAGATCCATCTCTTTTACCAAGTCAATCATCG 493

Qy 133 SerSerSerPheProSerProSerArgGlnProAsnAsnMetSerSerThPhe 152
 Db 492 TCTTCTTCAATTCCTTACGCTTCTCGGGGAC-----AACATC---TCCACTATC 445
 Qy 153 PheProPheLeuArgAngGlnGlyIleProSerSerLeuProSerLeuAlaGlnSerAsn 172
 Db 444 TTCCTTCTTCTCGAGAACGCTGGGATCCCTTCCTGCTCTCTCTCAGATATCTCAAC 385
 Qy 173 SerCyProValThrProProValSerSerProThSerValSerAsnProLeuPro 192
 Db 384 AGTCTCTCCGCTCCACCTGCTCTCTCCCAACTATATAACACCCCAAGCCCTTACT 325
 Qy 193 AsnTrpGlnSerIleAlaValGlnSerMetAlaIle---AlaValGlnSerMetAlaSer 211
 Db 324 ACTTGGAGGCTTTTACCAACAGTCAATGCGCATTCCTCTAAGCATGCAATGCTCTCT 265
 Qy 212 PheAsnTyrProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPhe 231
 Db 264 TTTAAGTACCGCTTCAACGGGTTTCTGCGCGGAGTCCCACTCATATCGCCAGTTC 205
 Qy 232 HisThrLeuAlaThrIleProGlnGlnAspGlnSerSerSerThrValAsnSerGly 251
 Db 204 AATGACCCGCTCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 145
 Qy 252 HisTrpIleSerPheGlnValPheAla---GlnGlnGlnProPheSer-----AlaSer 268
 Db 144 CATTTGATAGCTTTCAGAAATCTTCAACAGCAACAGCAACGCTTCCATGGGCTGCTCT 85
 Qy 269 MetValProThrSerProThrPheAsnLeuVally---ProAlaProGlnGlnMetSer 287
 Db 84 GCAGTGGCGGCTTCTCTTACTTCACTAGTGAACACACAGGCTCAGCGGCTATCG 25
 Qy 288 ProAsnThrAlaAlaPheGlnGlu 295
 Db 24 GCAAAACACTGCGGGAATCCAAAG 1
 RESULT 15
 AV822043 639 bp mRNA linear EST 01-APR-2002
 LOCUS AV822043 RAFL4 Arabidopsis thaliana cDNA clone RAFL04-20-E20 5',
 DEFINITION mRNA sequence.
 ACCESSION AV822043
 VERSION AV822043.1 GI:19864075
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eustosids II; Brassicales; Brassicaceae; Arabidopsie.
 1 (bases 1 to 639)
 Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
 Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
 Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
 and Shinozaki,K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b).
 TITLE
 JOURNAL
 COMMENT
 Unpublished
 Contact: Mocoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: meeki@rkc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). This clone is in a
 modified pBluescript vector as a SerI/XhoI insert. Please visit our
 web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
 further details.
 Location/Qualifiers
 1..639
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 20, 2003, 03:34:58 ; Search time 75 Seconds
(without alignments)
1977.396 Million cell updates/sec

Title: US-09-995-938A-7
Perfect score: 1765
Sequence: 1 MTSDDATSTSAIAAAAAAAAA.....IHDVGHDELTITGKANGK 336

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09995938.rn1 -MNMATCH=0.1 -LOOPEXT=0
-DB=Issued_Patents.NA -QPMF=fastap -SUFIX=p2n.rn1 -MINLEN=200000000
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMF=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09995938 @CGN 1.1 56 @runat 19122003 142607 4259 -NCPU=6 -ICPU=3
-NO MAP -LANG=ENGLISH -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA:
1: /cgn2_6/prodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	8.3	196	4	US-09-313-294A-6581
2	143	8.1	1926	4	US-09-249-585A-4
3	143	8.1	1931	2	US-09-130-114-2
4	141.5	8.0	12001	1	US-08-458-568A-11
5	140	7.9	9636	1	US-08-323-170B-1
6	140	7.9	9636	1	US-08-954-441-1
7	137.5	7.8	3132	2	US-08-224-482-3
8	137.5	7.8	3132	2	US-09-205-921-1
9	137.5	7.8	3132	4	US-09-300-958A-32
10	137.5	7.8	3132	4	US-09-702-705-320
11	137.5	7.8	3132	4	US-09-736-457-320
12	137.5	7.8	4401	4	US-09-614-034-192

13	135.5	7.7	1560	4	US-09-598-747-26
14	134.5	7.6	1209	1	US-08-314-309A-5
15	134.5	7.6	1513	1	US-08-314-309A-2
16	134.5	7.6	1546	1	US-08-314-309A-3
17	134.5	7.6	28958	1	US-08-258-261B-6
18	134.5	7.6	28958	1	US-08-456-837-6
19	134.5	7.6	28958	1	US-08-457-342-6
20	134.5	7.6	28958	1	US-08-457-646A-6
21	134.5	7.6	28958	1	US-08-458-076A-6
22	134.5	7.6	28958	1	US-08-764-233A-4
23	134.5	7.6	28958	1	US-08-457-335A-6
24	134.5	7.6	28958	1	US-08-729-214-6
25	134.5	7.6	28958	3	US-09-028-834-6
26	134.5	7.6	49377	1	US-08-764-233A-1
27	134.5	7.6	4403765	3	US-09-103-840A-2
28	134	7.6	2274	4	US-09-252-991A-3660
29	134	7.6	3297	4	US-09-252-991A-3615
30	133	7.5	4279	3	US-09-041-886-22
31	133	7.5	77536	4	US-09-410-551B-1
32	132	7.5	1853	1	US-08-553-110-2
33	132	7.5	4168	4	US-09-266-225D-17
34	130	7.4	2568	3	US-08-899-437-1
35	130	7.4	2568	3	US-09-126-121-1
36	130	7.4	4411529	3	US-09-103-840A-1
37	129.5	7.3	1681	4	US-09-434-288-5
38	129.5	7.3	4822	4	US-09-484-970B-106
39	128.5	7.3	1775	4	US-08-754-477A-1
40	128.5	7.3	1775	4	US-08-754-477A-4
41	128.5	7.3	1775	4	US-08-754-477A-120
42	128.5	7.3	1775	4	US-08-754-477A-124
43	127.5	7.2	816	4	US-08-754-477A-3
44	127.5	7.2	816	4	US-08-754-477A-122
45	127.5	7.2	816	4	US-08-754-477A-126

ALIGNMENTS

RESULT 1
US-09-313-294A-6581
Sequence 6581, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ico, Laura Y.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313, 294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 6581
LENGTH: 196
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700352049H1
LOCATION: 2, 134, 168, 174, 180, 183, 185, 189
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6581

Alignment Scores:

Pred. No.: 2.9e-05
Score: 147.00
Percent Similarity: 70.73%
Best Local Similarity: 68.29%
Query Match: 8.33%
DB: 4
US-09-995-938A-7 (1-336) x US-09-313-294A-6581 (1-196)

Length: 196
Matches: 28
Conservative: 1
Mismatch: 4
Indels: 8
Gaps: 1

```

Oy      50 GYLAATAGAGAAGlnclyasprTYAenLEUProLYHIScyasphsnaENGluValLeu 69
          |||
Db       5 GGCCCTCGGCGACACGGCGGTCAACAAGTCGCCAACACACTGCACAAACAAC----- 55.
Oy      70 LysAlaLeucyVALGlunlaGIYTpvAlvalGlunspGIYThrTYrArGLye 89
          |||||
Db       56 -----GAGCGCGGTGGTGTCTGTGAGCCCCGACCGCACCACTACCGCCAG 100
Oy      90 GIY 90
          |||
Db      101 GGA 103

RESULT 2
US-09-249-585A-4
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: HORLICK, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES
; FILE REFERENCE: 0867//00905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Alignment Scores:
Pred. No.:      0.00251      Length:      1926
Score:          143.00      Matches:     70
Percent Similarity: 40.89%    Conservative: 31
Best Local Similarity: 28.34%   Mismatches: 100
Query Match:      8.10%        Gaps:      47
DB:               4           Gaps:      8

```

[illegible]

Db 957 TCCTCCCGCTCTGCTCTTCACTCCGGGCCCACTCTCTCCGTCACTCCGGCCAGC 1016
 Qy 232 HistHrheua1aThrileProGluCyAspG1uSerAspSerSerThrValAspSerGly 251
 Db 1017 TCTCCATCACTCCGGGCC-----AGCTCTCATCACTCCGGC 1058
 Qy 252 HistP1eSerPheGlnLysPheAlaGlnGlnLmProPheSerAlaSerMetValPro 271
 Db 1059 GGGCCCATCTCTCGACT-----TCTCGGTCCCGCCCTTTCACACTTCTCGGTCCC 1112
 Qy 272 ThrSerProThrPheAsnLeuValLysProAlaProGln-----GlnMet 286
 Db 1113 CTCTCCAGCACCTG-CACCTCTTTCTCCGGGCTTCAAGGCTCATCACTAGTAGTAGTA 1171
 Qy 287 SerProAsnThrAlaAlaPheGlnLysGlnGlyGlnSerSerGluPheLysPheGluAsn 306
 Db 1172 GGCCCAAGGTGGGGCTCCGGGGAGAGT-C-----CAT 1204
 Qy 307 SerGlnValLysProTTPGlu 313
 Db 1205 CTTCGGTAAAGAGTGGAC 1225

```

RESULT 3
US-09-130-114-2
; Sequence 2, Application: US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassem B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130.114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

Alignment Scores:
Pred. No.: 0:00252 Length: 1931
Score: 143.00 Matches: 70
Percent Similarity: 40.89% Conservative: 31
Best Local Similarity: 28.34% Mismatches: 100
Query Match: 8.10% Indels: 47
Gaps: 8
2

US-09-995-938A-7 (1-336) x US-09-130-114-2 (1-1931)

QY 95 ProGlyluIleAlaGlyThrSerSerArgValThrProTyrSer---SerGlnAngln 113
Db 540 CCGGCTCCGCCGCTGCTGCTCTCCGCCGCTCCGCTCCGCTCCGCCGCTCTCTCC 599
QY 114 SerProIleSerSerAlaPheGlnSerProIleProSerGlyrGlnValSerProSer 133
Db 600 TCCCGGCTGCTGCTCTCCGCCGCTCCGCCGCTCCGCCGCTCTGCTCTCCGCCGCTCT 659
QY 134 SerSerPhePro-----SerProSerArgGlyGluPro 144
Db 660 TCCCTCTCCCGCTGCTCTCCGCCGCTCTCCGCCGCTGCTGCTCTCCGCCGCTCTCCG 719
QY 145 AsnAsnAsnMetSerSerThrPhePheProPheIleuArgAnglyGlyIlePro----- 162
Db 720 TCCGCTGCTCTCCGCCGCTCTGCTCTCCGCCGCTCCGCCGCTCTGCTCTCCGCCGCTCT 779
QY 163 -----SerSerLeuProSerLeuArgIleSer 171
Db 780 CCGTCTGCTCTCCGCCGCTCTCCGCCGCTCTCCGCCGCTCTCCGCCGCTCTCCGCCGCTCT 839

```



```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,441
FILING DATE: 20-OCT-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,170
FILING DATE: 13-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elnhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 015280-113110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 149..9556
US-08-954-441-1

```

Alignment Scores:	
Pred. No.:	0.0605
Score:	140.00
Percent Similarity:	37.60%
Best Local Similarity:	22.10%
Query Match:	7.93%
DB:	4
	9
Length:	9636
Matches:	68
Conservative:	23
Mismatches:	78
Indels:	74
Gaps:	9

US-09-995-938A-7 (1-336) X US-08-954-441-1 (1-9636)

Oy	102	SerSerArgValThrProTyrSerSerGlnAsnGlnSerProLeuSerSerAlaPheGln	121
		:::	
Db	1488	TCATCTCATATTCACCTCTTCTTCACCTACCTCTTACCTCTTCTTCACCTACCTCT	1429
Oy	122	SerProIleProSerTyrGlnValSerProSerSerSerSerPheProSerProSerArg	141
Db	1428	TCACCTACCCCTTCACCTACCTCTTCACCTTCTTTCACCTACCTCTTCACCTACCTCT	1365
Oy	142	GlyGlnProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAsnGlyGlyIle	161
		:::	
Db	1368	TCACCTACCTCTTCACCTACCTCTTCA-----	1342
Oy	162	ProSerSerLeuProSerLeuArgIleSerAsnSerCysProValThrProProValSer	181
		:::	
Db	1341	CCTACCTCTTCACCTACCTACCT-----TCTTCACCTACCTCTTCACCTACCTCT	129
Oy	182	SerProThrSerTyrAsnProIleAsnProLeu-----ProAsn	193
		:::	
Db	1296	TCACCTACCTCTTCATCTCATCTTCATTTTATTTAGTATCATCAAGATATCACCACTCT	1237
Oy	194	TyrGlnSer-----IleAlaLysGlnSerMetAlaIleAlaLysGln	207
Db	1236	TGGTCTTCAATATATTCGCGAACTTTATCATATCTCCATCCCGCTCTTACAGAAATCT	1177
Oy	208	SerMetAlaSerPheAsnTyrProPheTyrAlaValSerAlaProAlaSer-Pro-----	225
		:::	
Db	1176	TCATCTCATCTCATTAATAACTTCT---TCTGAAGATCTGACGACCTACTCTCTCGATGT	1120
Oy	226	-----ThrIleArgGlnGlnPheIleThr	233
Db	1119	TCTCTTGATATGTGTTCTTGCTTTCATCCCACTTCTTCATCAACATATCATCATATAT	1067

[illegible]

RESULT 7
 US-08-224-482-3
 : Sequence 3, Application US/08224482
 : Patent No. 5837692
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Mercola, Dan
 : APPLICANT: Adamson, Eileen D.
 : TITLE OF INVENTION: Inhibition of the Mitogenic Activity of
 : TITLE OF INVENTION: PDGF by Mammalian EGR
 : NUMBER OF SEQUENCES: 12
 : CORRESPONDENCE ADDRESS:
 :

ADDRESS: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA

COMPUTER READABLE FORM:

COMPUTER: IBM PC compati

SOFTWARE: PatentIn Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/224,482

CLASSIFICATION: 435

NAME: Campbell, Cathryn A.

REFERENCE/DOCKET NUMBER: P-ME 9913

TELEPHONE: (619) 535-9001

INFORMATION FOR SEQ ID NO: 3

LENGTH: 3132 base pairs

STRANDEDNESS: do

US-08-224-482-3

Alignment Scores

Score: _____

Best Local Similarity

DB:

US-09-995-938A-7 (1-336) X US-08-224-482-3 (1-3132)

QY 2 ThrSerAspGlyAlaThrSer-----AlaAlaAlaAla 15

	APPLICANT:	McClelland, Michael	
	APPLICANT:	Welsh, John	
	APPLICANT:	Trenkle, Thomas	
	TITLE OF INVENTION:	Reduced Complexity Nucleic Acid Targets and Methods of	
	TITLE OF INVENTION:	Using Same	
	FILE REFERENCE:	P-PH 3457	
	CURRENT APPLICATION NUMBER:	US/09/300,958A	
	CURRENT FILING DATE:	1999-04-27	
	PRIOR APPLICATION NUMBER:	60/083,331	
	PRIOR FILING DATE:	1998-04-27	
	PRIOR APPLICATION NUMBER:	60/098,070	
	PRIOR FILING DATE:	1998-08-27	
	PRIOR APPLICATION NUMBER:	60/118,624	
	NUMBER OF SEQ ID NOS:	85	
	SOFTWARE:	Patentin Ver. 2.0	
	SEQ ID NO 32		
	LENGTH:	3132	
	TYPE:	DNA	
	ORGANISM:	Homo sapiens	
	US-09-300-958A-32		
	Alignment Scores:		
	Pred. No.:	0.0171	Length: 3132
	Score:	137.50	Matches: 63.
	Percent Similarity:	41.06%	Conservative: 45
	Best Local Similarity:	23.95%	Mismatches: 92
	Query Match:	7.79%	Indels: 63
	DB:	Gaps: 4	Gaps: 9
Oy	US-09-995-938A-7 (1-336) x US-09-300-958A-32 (1-3132)		
Oy	2 ThisSerAspglyAlathrSertHiser-----Alalaalaaiala 15		
Db	1328 ACGAGCTCACCACGGCCATCGGCATCCACACAGGCGCACAGGCCATTCCAGTCCGCATCT		1387
Oy	16 AlalaalaaialaalAla-ArGaRgIvPProserTrPaRgluaRgluaanaAnaAr 35		
Db	1388 GCATTCGGCAACTTGCCGCGCAGCGCACCATCCACATCCGACCCACACAGCG 1447		
Oy	35 GaRGluARgaRGAraGalavalAlaalyeleryThrglyLeuatgalagngl 55		
Db	1448 AAAAGCCTTGCCGTGCGACAATCTGTGGAAAGAATTGCCAGAGCATGAACGCAGA 1507		
Oy	55 yASPTyrAnleuPrOlysHyScysASPANAangluValleubyalaLeuCysvalGl 75		
Db	1508 GGCATAACCAAGATCCATTGGGGCAGAAAGAACAAAAGTAGTTGT---- 1562		
Oy	75 uALeGIYTRPVaIVaIGluabPgLy-ThrtHtYarGuBgLYcylasProlaup 95		
Db	1563 -----GGCTCTTCGGCCACCTCTCTCTCTCTCTTCTTCTTACC 1597		
Oy	95 roGLygLuIlleaLAglyThrseseraRgvallThProTyTSerserglnan Ginserp 115		
Db	1598 Cgtcccgctggct-----AcctttTaccgcgtcccgcattTAacTCTTATCCatCCC 1651		
Oy	115 rOlEusSeSaLaPhegInseRProlliePROserTYrGlnValseRproSeSers 135		
Db	1652 CGGCGACACACCTATACCATTCCTCTGTGCCACACTCTTCTCTCCGCGCTCTCGA 1711		
Oy	135 erPheProSerPro--SerArglyGUprOAmaNaMaMeSeSerThrPhelep I54		
Db	1712 CCTACCCATCCCCCTGTGCAGATGGCTTCCCTCCCTCGGTGGCCACACAGTAC--- 1767		
Oy	154 ropHeLEuarGANdlyglyllePro-----SerSerleuPros 167		
Db	1768 -----TCTCTGTCTCCCTGCTTCCGCGCCACAGTCAAGCATTCCTT 1813		
Oy	167 etLeuarGileseranBercYSProVALthrProVaISerSerProTHiser----- 185		
Db	1814 CCTACGCTGTACCAAATCTTTCAGCGGCTCCACAGGGCTTGGACATGACAGCAACT 1873		
Oy	185 ----- 185		

```

Db      1874 TTCTCCAGACACATTGAATTGTTAAAGGGAAGGGGAAAAGAAAGGGAAGGAGGA 1933
Qy      186 --LysaenProlyseProleupProabentripglusertleialalyGlnSerMetAlaIle 205
Db      1934 AAAAGAAACACACAAAGACTTAAGAGACAGGAGG-----GAGAGAGATGGCCATTG 1981
Qy      205 latyGlnSerMetAlaSerPheantYrProPheTyAlaValaIleSerAlaProAlaSerP 225
Db      1982 GAGAGAGAGGGTCTCTTAGTCAATGAGAGTTCT-----TCAGAGCCAAAGTCTCTC 2032
Qy      225 rothr 226
Db      2033 CCTCT 2037

RESULT 10
US-09-702-705-320
; Sequence 320, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 320
; LENGTH: 3132
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-320

Alignment Scores:
Pred. NO.: 0.0171 Length: 3132
Score: 137.50 Matches: 63
Percent Similarity: 41.06% Mismatches: 45
Best Local Similarity: 23.95% Mismatches: 92
Query Match: 7.79% Indels: 63
DB: 4 Gaps: 9
US-09-995-938A-7 (1-336) x US-09-702-705-320 (1-3132)
Qy      2 ThreersepglyAlaThrSerThrSer-----AlaAlaAlaAlaAla 15
Db      1328 ACGAGCTCACCCGCCACATCCGACATCCACACAGGCCAGAAAGCCCTTCCAGTGGCCGACATCT 1387
Qy      16 AlalaAlaAlaAlaAla-ArgArgLyIsProSerTriArgGluuArgIuabenaenArgAr 35
Db      1388 GCATGCCGCACTThAGCCCGCAGCAGCACCACTCACACCACACATCCGACCCACACAGAGCG 1447
Qy      35 garGluuArgThrArgThrAlaValaAlaAlaAlaLyIleTyThrGlyLeuArgAlaGInGI 55
Db      1448 AAAAGCCCTTGGCTCGTGGACATCTGTGGAAGAAAGTTTCCAGAGAGCGATGAACGCAAGA 1507
Qy      55 yAspTyAsnLeuProLyshIsCySaapnaenAnGluValleuLySAlaLeuCySValGI 75
Db      1508 GGCATACCAGAAATCCACTTGGCGGAGAAAGCAAGAAAGACAGACAAAAGTGTGT----- 1562
Qy      75 uAlaGlyTrpValValGluGluuAbpGly-ThrThrTyraGlyGlyCyelysPProLeuP 95
Db      1563 -----GGCTCTTGGCGGACCACTCTCTCTCTCTTCCATAC 1597
Qy      95 roGlyGluIleAlaGlyThrSerSerArgValThrProTyrSerSerGlnaenGlnSerP 115

```

Db 1598 CGTCCCGGTTCT-----ACCTCTTACCGGCTCCCGGCTTACTACTCTTATCATCC 1651
Qy 115 rOlEusErSerAlaPhEgInSerProIleProSerTyGInValSerProSerSers 135
Db 1652 CGGACCAACCATCATACCATCCCTGTCACCACTCTTCTCTCCCGGCTCTCGA 1711
Qy 135 erPhEProSerPro---SerArgIyGluProAsnAsnAsnMetSerSerThrPhePheP 154
Db 1712 CCTACCCATCCCTGTGACAGTGGCTTCCCTCCCGGCTGGTGGCCACCATAC--- 1767
Qy 154 rOpHeLearGAsnIyGlyIlePro-----SerSerLeuProS 167
Db 1768 -----TCCTCTGTCCCTCCCTGCTTCCCGGCAAGGACAGCTTCCCTT 1813
Qy 167 erLeuArgIleSerAsnSerCySProValThrProProValSerSerProThrSer--- 185
Db 1814 CCTCAGCTGTACCAACTCTTCCAGGCGCTTCCAGAGGCTTTCGACATGACAGCAACT 1873
Qy 185 ----- 185
Db 1874 TTTCTCCAGCAATGAAATTCTCTAAGGAAAGGAAAGGAAAGGAAAGGAGA 1933
Qy 186 --LysAsnProIySerProLeuProAsnTrpGluSerIleAlaIyGInSerMetAlaIleA 205
Db 1934 AAAAGAAACACAGAGACTTAAGACAGAGAG-----GAGGAGATGGCCATAG 1981
Qy 205 lAlYsGInSerMetAlaSerPheAsnTyrrProPheTyAlaValSerAlaProAlaSerP 225
Db 1982 GAGAGAGAGGTTCTCTTAGTGACATGAGAGGTTT-----TCAGAGCCAAAGTCTTC 2032
Qy 225 roThr 226
Db 2033 CCTCT 2037

RESULT 11

US-09-736-457-320
Sequence 320, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Panger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Derrick
APPLICANT: Reiter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 320
LENGTH: 3132
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-320

Alignment Scores:

Pred. No.: 0.0171 Length: 3132
Score: 137.50 Matches: 63
Percent Similarity: 41.05% Conservative: 45
Best Local Similarity: 23.95% Mismatches: 92
Query Match: 7.79% Indels: 63
Gaps: 9

US-09-995-938a-7 (1-336) x US-09-736-457-320 (1-3132)

Qy 2 ThrSerAspGlyAlaThrSerThSer-----AlaAlaAlaAlaAla 15
Db 1328 ACAGCTCAACCGCCACATCCGATCCACAGAGCCAGAGCCCTTCCAGTCCGATCT 1387
Qy 16 AlaAlaAlaAlaAlaAla-ArgArgIyProSerThrPargIuArgIuAsnAsnArg 35
Db 1388 GCATCGCACTTCAGCGCGGACGACACCTACACCAACATCCGACCCACACAGGCC 1447
Qy 35 gArgIuArgIuArgArgAlaValAlaAlaIyAlaIyTyrrThrGlyLeuArgAlaIngl 55
Db 1448 AAAAGCCCTTCGCTGACATCTGTGAAAGAAATTGCCAGAGCGATGAAGCAGAG 1507
Qy 55 yAsPTrAsnLeuProIyShISCyAsPAsnAsnIyValLeuValAlaLeuCyAlaIeI 75
Db 1508 GGCATACCAAGATCCATCTGCGGAGAGAGCAAGCAAGAGCAAGAAAGTGTGT----- 1562
Qy 75 uAlaIyTrpValValIuGluIuAspGly-ThrThyArgIyGlyCySlyProLeuP 95
Db 1563 -----GACCTCTGCGCACCTCTCTCTCTCTCTCTTCTTCTTACC 1597
Qy 95 rOgIyGluIleAlaIyThrSerSerArgValThrProTySerSerGlnAsnGlnSerP 115
Db 1598 CGTCCCGGTTGCT-----ACCTCTTACCGGCTCCCGGTTACTACTCTTATCCATCCC 1651
Qy 115 rOlEusErSerAlaPhEgInSerProIleProSerTyGInValSerProSerSers 135
Db 1652 CGGACCAACCATCATACCATCCCTGTGACCACTCTCTTCTTCCCTCCGCTCTCGA 1711
Qy 135 erPhEProSerPro---SerArgIyGluProAsnAsnAsnMetSerSerThrPhePheP 154
Db 1712 CCTACCCATCCCTGTGACAGTGGCTTCCCTCCCGGCTGGCGACACCATAC--- 1767
Qy 154 rOpHeLearGAsnIyGlyIlePro-----SerSerLeuProS 167
Db 1768 -----TCCTCTGTCCCTCCCTGCTTTCGAGCCAGGTCAGACACTTCCCTT 1813
Qy 167 erLeuArgIleSerAsnSerCySProValThrProProValSerSerProThrSer--- 185
Db 1814 CCTCAGCTGTACCAACTCTTCCAGGCGCTTCCAGAGGCTTTCGACATGACAGCAACT 1873
Qy 185 ----- 185
Db 1874 TTTCTCCAGCAATGAAATTCTCTAAGGAAAGGAAAGGAAAGGAAAGGAGA 1933
Qy 186 --LysAsnProIySerProLeuProAsnTrpGluSerIleAlaIyGInSerMetAlaIleA 205
Db 1934 AAAAGAAACACAGAGACTTAAGACAGAGAG-----GAGGAGATGGCCATAG 1981
Qy 205 lAlYsGInSerMetAlaSerPheAsnTyrrProPheTyAlaValSerAlaProAlaSerP 225
Db 1982 GAGAGAGAGGTTCTCTTAGTGACATGAGAGGTTT-----TCAGAGCCAAAGTCTTC 2032
Qy 225 roThr 226
Db 2033 CCTCT 2037

RESULT 12

US-09-614-034-192
Sequence 192, Application US/09614034
Patent No. 6489307
GENERAL INFORMATION:
APPLICANT: PHILLIPS, M. IAN
APPLICANT: ZHANG, YUAN
TITLE OF INVENTION: ANTISENSE COMPOSITIONS TARGETED TO BETAL-ADRENOCEPTOR-SPECIFIC
FILE REFERENCE: METHODS OF USE
CURRENT APPLICATION NUMBER: US/09/614,034
CURRENT FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 09/152,717
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US99/21007
PRIOR FILING DATE: 1998-09-14
NUMBER OF SEQ ID NOS: 204

Oy 87TyrAglYglIcYueYbPrOleuPrOglYglUlleAlaGlIhRSeSaRgVAlThr 106
 545 TCGCGCGCGGGGGCCAGCTCACCACCAACCCAGCTCGAAGACT--TCCGGGGTTCC 601
 Oy 107 PrOTYrSerSerGlNaAnGInSerPrOleuSerSerAlaPheGInSerProIleProSer 126
 Db 602 CCGAGGGGAGATCCTCGCGGGGAGACTCATGGAATGGGTGCGCGCGGCACTCCTCCGTTCC 661
 Oy 127 TyrGInValSerPrOSeSerSerSerPheProSerProSerAlrgIgluPrOaAn 146
 Db 662 -----GCACACACATCATCTCCGAGACCGTCAACCGCGGTCACTTCTCCG----- 706
 Oy 147 AAmMetSerSerThrPhePheProPheLeuArgAnGlyIleProSerSerLeuPro 166
 Db 707 -----CCCGCCCTT 715
 Oy 167 SerLeuArgIleSerAnSerCySProValThrProProValSerSerProThr----- 184
 Db 716 TCC-----GCGTGCCTCGAGCTCCACACCGGTGCTGCCGAGCGCGTCCG 760
 Oy 185 -----SerLyAnbProLybProleuProaAntTp 194
 Db 761 TCGTGCACCGCGCGCGCTCGCGCGCGAGACTCCACTTCGCGGCTCGAGACGCTTACTGGA 820
 Oy 195 GluSerIleAlaIleGInSerMetAlaIleAlaIleGInSerMetAlaSerPheAntYr 214
 Db 821 ACCGGGGGATCTCAGCTTCGCGCGTGGCGAGCGGGCGCGCCCAATCTTCAGAACAAAC 880
 Oy 215 ProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGInPheHisThrLeu 234
 Db 881 CCATGCGCGCTCA---TCGGCGGGCGGAGCTCCGCCATGAGAGAGTCCAACTTCTTCAACA 937
 Oy 235 AlaThrIleProGluCySPaPgluSerPaPSeSerThr 247
 Db 938 AGTACGGCTCCCATGTGTATCATCATCAGCCGCGCAACA 976
 RESULT 14
 US-08-314-309A-5
 Sequence 5, Application US/08314309A
 Patent No. 5677141
 GENERAL INFORMATION:
 APPLICANT: ISOGAI, TAKAO
 APPLICANT: FUKAGAWA, MASAO
 APPLICANT: IWAMI, MORITR
 APPLICANT: ARAMORI, ICHIRO
 APPLICANT: KOJO, HITOSHI
 TITLE OF INVENTION: PROCESS FOR PRODUCING 7-AMINOCEPHEN
 TITLE OF INVENTION: COMPOUND OR SALTS THEREOF
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUBSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/314,309A
 FILING DATE: 30-SEP-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/631,906
 FILING DATE: 21-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Oblon, No. 5677141man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 18-863-0 CONT

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAN UR
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1209 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1206
US-08-314-309A-5

Alignment Scores:
Pred. No.: 0.00701 Length: 1209
Score: 134.50 Matches: 87
Percent Similarity: 33.15% Conservative: 33
Best Local Similarity: 24.03% Mismatches: 122
Query Match: 7.62% Indels: 121
DB: 1 Gaps: 14

US-09-995-938A-7 (1-336) x US-08-314-309A-5 (1-1209)
QY 2 ThrSerAspGly---AlaThrSerThSerAlaAlaAlaAlaAlaAlaAlaAla 20
Db 140 ACATCTCTGGGTTCAGCGACATCCACAAAGCGAGCTCAGCCGACAGCCCGGTA 199
QY 21 AlaArgArgLysProSer----- 26
Db 200 TCGAGAAAGAGTCCACATCGACACCTTTAACGCTATGTCGGCGAGTTGACGAGACTA 255
QY 27 ---TTPaRGGLuArgGLuAsnAspArgArgArgGLuArgArgArgAlaAla-Al 45
Db 260 CCATTGAGAGAGATCAAGAACAAACCGGATGTCCTCGAGGTAGAGGAGACCAAGATCGGC 319
QY 45 alysllelTyThRgLyLeuArgAlaGlnGlyAspTyRAsnLeuProLysHieCyAspAs 65
Db 320 ACCTCTTGACAGAGGAGAGACGAGGAGAGAATTGACGACCGCCGCTCTGACCCAGAACG 379
QY 65 nasngluValleuylsAlaLeuCyAsvalGluAlaGlyTTPvalAlaGluGluAspGlyTh 85
Db 380 GCGCCTGGGGGCTGGGCGACCATCTCTCATTCGCGACAGCTGGCTGCAGCACACT-ACATCTAC 438
QY 85 rThTyTrArgLyAsGlyCyLysProLeuProGlyLuiAlaGlyThSerSerArgva 105
Db 439 GACGACTGCGCGGCGAGCGGACCTACGCTGAGTGGAGACCGGCGCATCTCTGAGAGT 498
QY 105 lThProTySerSerGlnAsnGlnSerProLeuSerSerAlaPheGlnSerProIlePr 125
Db 499 CACA-----ACGAGTTCTCCGGCCG- 11
QY 125 oSerTyGlnValSerProSerSerSerSerPheProSerProSerArgGlyLuiProAs 145
Db 523 ATCAGCGGCTACAAACGCCGCTGGCGGAGAACAGCGGACACCAACGCGCACCGCACCCAC 582
QY 145 nasn-----AsnMeSerSerThrPhehePro-----Ph 155
Db 583 GTCTGTGGCACCATTTGGCGGACGAGCACTTACGGCGTTGCCAAGAACCAACCTCATCGCT 642
QY 155 eleuArgAsnGlyGly-----ileProSerSerLeuProSer----- 167
Db 643 GTCAAGGCTCTCGGGGAGATCTTCAGCTCTACTTCTCATATCTCTTGAACGCTTCAACTGG 702
QY 168 -----LauArgIleSerAsnSerCyAsPro-- 175
Db 703 GCCGTGAACGATATCATCAACAGGGGCGCCAGAACAAAGGCTGCCATCGACAGATGCTCTG 762
QY 176 -----ValThProProValse 181
Db 763 GGATGGCTACTCTTCTTGCTTCAACAACGCCGCTAAACACTGCTTACTCTCCGCGCGCTC 822

```

```

Oy 101 rSerProthr-----SerLyAsnProLyAspProle 191
Db 823 CTCTCGGTGTCGCGCCGCGGACGATTAACGAGACCGCGCACTACTCCCGCGCTCG 882
Oy 191 uProAsnTrpGluSerIleAlaIleGlnSerMetAlaIleAlaIleGlnSerMetAla 211
Db 883 GCGC----- 886
Oy 211 rPheAsnTrpProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPh 231
Db 887 -----CCAGCCGCGCATCACCGGCTCGCTCATCGCTCCAACT-----GGCGCCGC 930
Oy 231 eHisThrIleuAlaThrIleIleProGluCyAspArgLysSerAspSerSerThrValAspSerI 251
Db 931 TCACACTTCAGCACTACCGGCTCGCTCGCTCGACATCT----- 967
Oy 251 yHisTrpIleSerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSerMetValPr 271
Db 968 -----TCGCCCCCGGAAACGACATCTCTCGCTCGATCGCGGACGACCTCGGC 1017
Oy 271 oHisSerProThrPheAsnLeuValLysPheProAlaProGlnGlnMetSerProAsnThrAl 291
Db 1018 A-----CCAACACCATCTCCGCGACCTCATGGCCACCCCATCTCATCGCGCTCGTC 1071
Oy 291 aAla 292
Db 1072 CTCT 1075

```

RESULT 15

US-08-314-309A-2
Patent No. 5677141
Sequence 2, Application US/08314309A

GENERAL INFORMATION:

APPLICANT: ISOGAI, TAKAO
APPLICANT: FUKAGAWA, MASAO
APPLICANT: IWAMI, MORITA
APPLICANT: ARAMORI, ICHIRO
APPLICANT: KONO, HITOSHI
TITLE OF INVENTION: PROCESS FOR PRODUCING 7-AMINOCEPHEM
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/314,309A
FILING DATE: 30-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/631,906
FILING DATE: 21-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5677141man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-863-0 CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1513 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
US-08-314-309A-2
Alignment Scores:
Pred. No.: 0.01
Score: 134.50
Percent Similarity: 33.15%
Best Local Similarity: 24.03%
Query Match: 7.62%
Db: 1
Gaps: 14
Length: 1513
Matches: 87
Conservative: 33
Mismatches: 132
Indels: 121
Gaps: 14

```

US-09-995-938A-7 (1-336) x US-08-314-309A-2 (1-1513)

```

Oy 2 ThisSerArgLy-----AlaThrSerThrAlaAlaAlaAlaAlaAlaAla 20
Db 229 ACATCTCTGGGTGTCGCGCATCTCAACAGCGAGCTTCAGCGCGCGGCGGTA 288
Oy 21 AlaArgLyAspProSer----- 26
Db 229 TCGAGAGAGAGTTCACATTCGACCTTTAAGCCTATGTCGCGGAGTTCGAGACATTA 348
Oy 27 ---TTPATGGLuArgGluAsnAsnArgArgArgGluArgArgArgAlaAlaAlaAla 45
Db 349 CCATTGAGGAGATCAAGAAACCGGAGATTCCTCGAGGTAGAGAGACAGATCTGGC 408
Oy 45 alysiLetyThrArgLysAlaGlnGlnLysAspTyrAsnLeuProLyHisCyAspAsp 65
Db 409 ACCTCTTCGAGAGAGAGAGAGAGAGAGAGATTCAGACCGCCCTCTCACCACGAGACG 468
Oy 65 nAsnGluValLeuLysAlaLeuCyValGluValGluValGluValGluValGluVal 85
Db 469 GCGCTGGGCGCTGGGACCATCTCTACCGACCTGCTGACCTGACCTACATCTAC 527
Oy 85 rThrTyrArgLysGlyCyLysProLeuProGlyLysIleAlaGlyThrSerSerArgVa 105
Db 528 GACGACTGCGCGGACGCGGACCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 587
Oy 105 lThrProTyrSerSerGlnAsnGlnSerProLeuSerSerAlaPheGlnSerProIlePr 125
Db 588 CACA-----ACGAGTCTCCGCGCCG-----CC 611
Oy 125 cSerTyrGlnValSerProSerSerSerSerPheProSerProSerArgGlyGluProAs 145
Db 612 ATCAGGCGCTACACGCGGCTGCGGAGCAAGCCGACACCAACGCGCACCGCACCCAC 671
Oy 145 nAsn-----AsnMetSerSerThrPhePhePro-----Ph 155
Db 672 GTGCGTGACACCATTCGCGGACGACGACGCGGCTTGCACAAACCAACCACTCATCGCT 731
Oy 155 eLeuArgAsnGlyI-----IleProSerSerLeuProSer----- 167
Db 732 GTCAAGGCTTCGCGGAGATCTTCGAGCTTACTTCATCATCTCTGACGCGCTCAACTGG 791
Oy 168 -----LeuArgLiesSerAsnSerCyPro-- 175
Db 792 GCGGTGAACGATCATCATCAACAGGCGCGGCAAGAAACAGGCTGCATGCTCCG 851
Oy 176 -----ValThrProProValSe 181
Db 852 GGTGTGCTACTCTTTCGCTTCAACAGCGCGTCAACACTGCTACCTCCGCGCGCTC 911
Oy 181 rSerProThr-----SerLyAsnProLyAspProle 191
Db 912 CTCTCGGTGTCGCGCGGACGATTAACGAGACCGCGCACTACTCCCGCGCTCG 971
Oy 191 uProAsnTrpGluSerIleAlaIleGlnSerMetAlaIleAlaIleGlnSerMetAla 211
Db 972 GCGC----- 975
Oy 211 rPheAsnTrpProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPh 231

```


1	2687	100.0	2687	13	US-09-995-9384-2	Sequence 2, Appl
2	2685.4	99.9	2687	13	US-09-995-9384-1	Sequence 1, Appl
3	1102	41.0	1456	13	US-09-995-9384-5	Sequence 5, Appl
4	1100.4	41.0	1456	13	US-09-995-9384-9	Sequence 4, Appl
5	713	26.5	3000	13	US-09-995-9384-3	Sequence 9, Appl
6	711.4	26.5	3000	13	US-09-995-9384-3	Sequence 3, Appl
7	535.6	19.9	1213	13	US-09-995-9384-12	Sequence 12, Appl
8	534	19.9	1213	13	US-09-995-9384-11	Sequence 11, Appl
9	149	5.5	831	10	US-09-978-8424-385	Sequence 3085, Ap
10	121.4	4.5	413	10	US-09-978-574-5002	Sequence 5002, Ap
11	109.6	4.1	284	10	US-09-878-574-14631	Sequence 14631, A
12	71.4	2.7	3673778	13	US-10-312-84-1	Sequence 1, Appl
13	70	2.6	6145	13	US-10-311-655-945	Sequence 945, Appl
14	63.2	2.4	3673778	13	US-10-312-84-2	Sequence 2, Appl
15	62.6	2.3	9642	13	US-10-311-655-330	Sequence 330, Ap

C	16	61	2.3	3673778	13	US-10-311-4841-1	Sequence 1, Appl1
	17	60.8	2.3	1603	13	US-10-311-455-1377	Sequence 1377, Ap
	18	60.6	2.3	5678	13	US-10-311-455-1111	Sequence 1111, Ap
	19	60.4	2.2	6070	13	US-10-311-455-1652	Sequence 1652, Ap
	20	60.4	2.2	6070	13	US-10-340-485-132	Sequence 132, App
	21	60.4	2.2	6954	13	US-10-311-455-1364	Sequence 1364, App
	22	60.4	2.2	14006	13	US-10-311-455-1991	Sequence 1931, Ap
C	23	60.2	2.2	6283	13	US-10-311-455-61	Sequence 61, Appl
	24	60.2	2.2	113515	13	US-10-311-455-2147	Sequence 2147, Ap
	25	60	2.2	15373	13	US-10-311-455-439	Sequence 439, App
	26	59.6	2.2	5649	13	US-10-311-455-822	Sequence 822, App
	27	59.6	2.2	16891	13	US-10-311-455-626	Sequence 626, App
	28	59.6	2.2	16891	13	US-10-340-485-58	Sequence 58, Appl
C	29	59.4	2.2	529	10	US-09-983-965-2109	Sequence 2109, Ap
	30	59.4	2.2	20486	13	US-10-340-485-164	Sequence 164, App
	31	59.2	2.2	4412	13	US-10-311-455-460	Sequence 462, App
	32	59	2.2	616	10	US-09-960-3352-3400	Sequence 3400, Ap
	33	59	2.2	5195	13	US-10-311-455-894	Sequence 894, App
	34	59	2.2	6351	13	US-10-311-455-1720	Sequence 1720, Ap
	35	59	2.2	6944	15	US-10-172-086-112	Sequence 112, App
	36	59	2.2	9539	13	US-10-340-453-54	Sequence 54, Appl
	37	59	2.2	9539	15	US-10-239-676-52	Sequence 52, Appl
C	38	58.4	2.2	6061	13	US-10-311-455-113	Sequence 113, App
	39	58.4	2.2	6189	13	US-10-340-485-146	Sequence 146, App
	40	58.4	2.2	6668	13	US-10-311-455-1668	Sequence 1668, Ap
	41	58.4	2.2	6668	13	US-10-311-455-1670	Sequence 1670, Ap
	42	58.4	2.2	17721	13	US-10-311-455-1701	Sequence 1701, Ap
	43	58.2	2.2	7490	13	US-10-311-455-255	Sequence 255, App
	44	58.2	2.2	11944	13	US-10-311-455-2159	Sequence 2159, App
	45	58	2.2	12705	13	US-10-311-455-121	Sequence 121, App

ALIGNMENTS

RESULT 1
ME-09-006-0788-2

```

Sequence 2, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
TITLE OF INVENTION: HORMONE ACTION IN PLANTS
FILE REFERENCE: SALKINS 046A
CURRENT APPLICATION NUMBER: US/09/995,938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 2687
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-2

```

Query Match	100.0%;	Score 2687;	DB 13;	Length 2687;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2687; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY	1	ACTTTAGTGTTTGGCTAATTCATCGAACCCCTGTGATTCATTCGAAATGTTTCCCACTGGC	60
Db	1	ACTTTAGTGTTTGGCTAATTCATCGAACCCCTGTGATTCATTCGAAATGTTTCCCACTGGC	60
QY	61	GTTGATGTTGGGGTTCCTCGCTTTTAAACATAGTATACATGCTCGTCACTTTCTAC	120
Db	61	GTTGATGTTGGGGTTCCTCGCTTTTAAACATAGTATACATGCTCGTCACTTTCTAC	120
QY	121	AAATTAACGTCATATTCACAAAATATTATACATGATTCATACATATATGCGCGGAACC	180
Db	121	AAATTAACGTCATATTCACAAAATATTATACATGATTCATACATATATGCGCGGAACC	180
QY	181	TTGTTATGGGCAAACTGTAACCCCTTTTCCCTTATGTTCAATGAATCATACAAGT	240
Db	181	TTGTTATGGGCAAACTGTAACCCCTTTTCCCTTATGTTCAATGAATCATACAAGT	240

Db 181 TTGTTATGGGACAACTGTAACCCCTTTTTCCTTTATGTCAATGAACTATCAAGT 240
Qy 241 TTTGGTTATGAATACATAAATTAATGATGAGACCCAGCAATTAATCCAAAATTGATATTA 300
Db 241 TTTGGTTATGAATACATAAATTAATGATGAGACCCAGCAATTAATCCAAAATTGATATTA 300
Qy 301 GATATCAAGCTTAATAATCAACATGTAACCAAACTTAATCTTTATAGAACATAGTAAT 360
Db 301 GATATCAAGCTTAATAATCAACATGTAACCAAACTTAATCTTTATAGAACATAGTAAT 360
Qy 361 GGTATTCACCAATCTTTATATCATTTGTAGGTAGGAGAGAGTAAAGAGAGAGC 420
Db 361 GGTATTCACCAATCTTTATATCATTTGTAGGTAGGAGAGTAAAGAGAGAGC 420
Qy 421 CAGGTATACATCAACATAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 421 CAGGTATACATCAACATAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 481 TATGCAAG 540
Db 481 TATGCAAG 540
Qy 541 CTTACCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
Db 541 CTTACCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
Qy 601 AGTAAGGAG 660
Db 601 AGTAAGGAG 660
Qy 661 TTTATTTGGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
Db 661 TTTATTTGGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
Qy 721 ACACGTTACTAAG 780
Db 721 ACACGTTACTAAG 780
Qy 781 CGTAACTCAACCTTCCAAAACCATGTAATTAACAGTGGAGAGAGAGAGAGAGAGAG 840
Db 781 CGTAACTCAACCTTCCAAAACCATGTAATTAACAGTGGAGAGAGAGAGAGAGAGAG 840
Qy 841 TGATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
Db 841 TGATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
Qy 901 TGAAG 960
Db 901 TGAAG 960
Qy 961 GAAAG 1020
Db 961 GAAAG 1020
Qy 1021 CAGTCTACGTTACCAATCTTTCAACCACTTAATTAAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 CAGTCTACGTTACCAATCTTTCAACCACTTAATTAAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 GGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGG 1140
Db 1081 GGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGG 1140
Qy 1141 AGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
Db 1141 AGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
Qy 1201 TCGAG 1260
Db 1201 TCGAG 1260
Qy 1261 TCAAGGTATTAATTTGGCTTAACATTTGTGATTAATTAAGAGAGAGAGAGAGAG 1320
Db 1261 TCAAGGTATTAATTTGGCTTAACATTTGTGATTAATTAAGAGAGAGAGAGAGAG 1320

Qy 1321 TGTGAAGCTGTTGGTTGGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 TGTGAAGCTGTTGGTTGGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 CTGCAATTTTTCAG 1440
Db 1381 CTGCAATTTTTCAG 1440
Qy 1441 CGTTGATTTCAATTTGAG 1500
Db 1441 CGTTGATTTCAATTTGAG 1500
Qy 1501 TGAATTTGAATCATCAG 1560
Db 1501 TGAATTTGAATCATCAG 1560
Qy 1561 TTTCTTTCAATTTTGAAG 1620
Db 1561 TTTCTTTCAATTTTGAAG 1620
Qy 1621 CTTGCTCCAAAGTCATGAG 1680
Db 1621 CTTGCTCCAAAGTCATGAG 1680
Qy 1681 TTTGGTTATTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTT 1740
Db 1681 TTTGGTTATTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTT 1740
Qy 1741 GCTGGAGCTTATCTCGAGTAATCTCCATATTCATTAACGAAACAGAGAGAGAGAGAG 1800
Db 1741 GCTGGAGCTTATCTCGAGTAATCTCCATATTCATTAACGAAACAGAGAGAGAGAGAG 1800
Qy 1801 GCTTTCAAGTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 1860
Db 1801 GCTTTCAAGTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 1860
Qy 1861 CTTTCTCGGAGTGAACCAATTAACCAATGATGCTTCAATCTTCCCTTCCAGAAAT 1920
Db 1861 CTTTCTCGGAGTGAACCAATTAACCAATGATGCTTCAATCTTCCCTTCCAGAAAT 1920
Qy 1921 GGTGGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1980
Db 1921 GGTGGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1980
Qy 1981 CCGGTCATGCGGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
Db 1981 CCGGTCATGCGGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
Qy 2041 AAGCAATCATGGCCATTTGCTAAACCAATCAATGAGAGAGAGAGAGAGAGAGAG 2100
Db 2041 AAGCAATCATGGCCATTTGCTAAACCAATCAATGAGAGAGAGAGAGAGAGAGAG 2100
Qy 2101 GTTTCTGCACTGCTAGTCCGACACATGCGCACAGAGAGAGAGAGAGAGAGAGAG 2160
Db 2101 GTTTCTGCACTGCTAGTCCGACACATGCGCACAGAGAGAGAGAGAGAGAGAGAG 2160
Qy 2161 GAAATGTAATGAG 2220
Db 2161 GAAATGTAATGAG 2220
Qy 2221 TTTGCAACAACAAGCCATTTCTGCTCTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
Db 2221 TTTGCAACAACAAGCCATTTCTGCTCTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
Qy 2281 GTGAAACCTGGGCTCAGACAGATGCTCAAAATATGCTGCTTCCAGAGAGAGAGAGAG 2340
Db 2281 GTGAAACCTGGGCTCAGACAGATGCTCAAAATATGCTGCTTCCAGAGAGAGAGAGAG 2340
Qy 2341 AGCTCTGAGTTAATTTGAAGATAGCCAGTTAAACCTGGAGAGAGAGAGAGATACAT 2400
Db 2341 AGCTCTGAGTTAATTTGAAGATAGCCAGTTAATACCTGGAGAGAGAGAGATACAT 2400

QY 2401 GATGGGTATGAGATCTTGAGCTTACCTTGGAAATGGAGGCGTGGTTCAT 2460
DB 2401 GATGGGTATGAGATCTTGAGCTTACCTTGGAAATGGAGGCGTGGTTCAT 2460
QY 2461 AAACAACTAGGCAAAACCCAAATGSCATGTCATTTGAAATATGAGAACTATCTCTTGAG 2520
DB 2461 AAACAACTAGGCAAAACCCAAATGSCATGTCATTTGAAATATGAGAACTATCTCTTGAG 2520
QY 2521 TATTTCTCTCTGTCAGAGTATTTGATCTTTATGGAATCTATGATGTTCTTCACTTAT 2580
DB 2521 TATTTCTCTCTGTCAGAGTATTTGATCTTTATGGAATCTATGATGTTCTTCACTTAT 2580
QY 2581 TATCCAAATATGCTGCCAAAGCTTCTCCATGGAAGATGAGTGTAGAGTAT 2640
DB 2581 TATCCAAATATGCTGCCAAAGCTTCTCCATGGAAGATGAGTGTAGAGTAT 2640
QY 2641 TCATTCATGAAATTTGGTTTCAAAAGCATTTATTTGATATAAAAA 2687
DB 2641 TCATTCATGAAATTTGGTTTCAAAAGCATTTATTTGATATAAAAA 2687

RESULT 2

US-09-995-938a-1
Sequence 1, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS.046A
CURRENT APPLICATION NUMBER: US/09/995.938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2687
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-1

Query Match 99.9%; Score 2685.4; DB 13; Length 2687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2686; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTTAGTTGGCCATTCATGAAACCTGATTCATTCCTCAATGTTCCCACTGC 60
DB 1 ACTTAGTTGGCCATTCATGAAACCTGATTCATTCCTCAATGTTCCCACTGC 60
QY 61 GTTGATGTTGGGTTCCCTGCTTTTAAGACATTAGTTCATGCTCTGCACTTTCTAC 120
DB 61 GTTGATGTTGGGTTCCCTGCTTTTAAGACATTAGTTCATGCTCTGCACTTTCTAC 120
QY 121 AAAATAAAGCTATATCCAAAAATTTATGATGATCATCATATATATGCGCGGAAC 180
DB 121 AAAATAAAGCTATATCCAAAAATTTATGATGATCATCATATATATGCGCGGAAC 180
QY 181 TTTGTAAGGCAAACTGTAACCCCTTTTCTTTTATGTTGATGATGATGATGAT 240
DB 181 TTTGTAAGGCAAACTGTAACCCCTTTTCTTTTATGTTGATGATGATGATGAT 240
QY 241 TTTGTTATGAT 300
DB 241 TTTGTTATGAT 300
QY 301 GATACCTAAAGCTTAAATCAATGTAACCAACTAAATATCTTTATGAAATAGTAAT 360
DB 301 GATACCTAAAGCTTAAATCAATGTAACCAACTAAATATCTTTATGAAATAGTAAT 360
QY 361 GGTATTCACCAATCTTATATATATATATATATATATATATATATATATATAT 420
DB 361 GGTATTCACCAATCTTATATATATATATATATATATATATATATATATATAT 420

QY 421 CAGTGTACATACACTATATACAGACAAATAGTCAAAATGTTCTTAAAGTGAATTTG 480
DB 421 CAGTGTACATACACTATATACAGACAAATAGTCAAAATGTTCTTAAAGTGAATTTG 480
QY 481 TATGCAAGAAAAGTAT 540
DB 481 TATGCAAGAAAAGTAT 540
QY 541 CTTCACCTGTAT 600
DB 541 CTTCACCTGTAT 600
QY 601 AGTAAGGAGCAACACTGAGAAAGCGGTCCCACTGTGTCTCTCTCTCTCTCTCT 660
DB 601 AGTAAGGAGCAACACTGAGAAAGCGGTCCCACTGTGTCTCTCTCTCTCTCTCT 660
QY 661 TTTATTTGGTAT 720
DB 661 TTTATTTGGTAT 720
QY 721 ACAGCTTACTATAGGTTATATATATATATATATATATATATATATATATATATAT 780
DB 721 ACAGCTTACTATAGGTTATATATATATATATATATATATATATATATATATATAT 780
QY 781 CGTATCTACCTCTTCCAAACATGTAATTTTACATGTAAGATGATGATACACAAACA 840
DB 781 CGTATCTACCTCTTCCAAACATGTAATTTTACATGTAAGATGATGATACACAAACA 840
QY 841 TGATTTCTTAATATGAT 900
DB 841 TGATTTCTTAATATGAT 900
QY 901 TGAATAAAT 960
DB 901 TGAATAAAT 960
QY 961 GAAGAAAGGCTATCTCTGTAAGCACTTCTCACTCTCTCTCTCTCTCTCTCTCTCAT 1020
DB 961 GAAGAAAGGCTATCTCTGTAAGCACTTCTCACTCTCTCTCTCTCTCTCTCTCTCAT 1020
QY 1021 CAGTGTACATACACTATATACAGACAAATAGTCAAAATGTTCTTAAAGTGAATTTG 1080
DB 1021 CAGTGTACATACACTATATACAGACAAATAGTCAAAATGTTCTTAAAGTGAATTTG 1080
QY 1081 GGGTTGTTGTTGTTTCCGATGATCTTGGATGAGCTATGACATTCAGACCTGTC 1140
DB 1081 GGGTTGTTGTTGTTTCCGATGATCTTGGATGAGCTATGACATTCAGACCTGTC 1140
QY 1141 AGCTGCGGCGGCGAGAGCGGCGAGAGAGAGCGTCTGAGAGAAAGGAGATTA 1200
DB 1141 AGCTGCGGCGGCGAGAGCGGCGAGAGAGAGCGTCTGAGAGAAAGGAGATTA 1200
QY 1201 TCGAGAGAGAAAG 1260
DB 1201 TCGAGAGAGAAAG 1260
QY 1261 TCAAGGTAT 1320
DB 1261 TCAAGGTAT 1320
QY 1321 TGTTAAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1380
DB 1321 TGTTAAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1380
QY 1381 CTTCATTTTTCACATCTGAGCTTCTTTATATATATATATATATATATATATAT 1440
DB 1381 CTTCATTTTTCACATCTGAGCTTCTTTATATATATATATATATATATATATAT 1440
QY 1441 CGTTGATTTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1500
DB 1441 CGTTGATTTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1500
QY 1501 TGAATTTGAATATACAGAACTATATGATGATCTATTTCTCAAACTGAATTTAT 1560

QY	1611	ATGTATGATACCTTGGTCCAAAGTCATGTCATGTTGTTCTTTGGCTTAACCGTGAATGAT	1670
Db	1845	ATTTA-----GATCTCTCTTAAGCTTTGATTTGA	1874
QY	1671	TGATGTTGATGGTTATATCTTTTGCTTTGGTGGAGTACAGGAGTGCACAGCTTTACC	1730
Db	1875	TTGTGACATTGTGATGTTTTTGGTATATGTTCTTATAGAGACACAGGAGACAAAGCTTACC	1934
QY	1731	TGTTGAGATAGCTGGGACTTCATCTCGAGTAACTTCATATTCATACAGAAACGAGGCC	1790
Db	1935	TGGTGACATGGCTGGATCATCTTCGAGCAACTCTTACCTTCCCTAACCAAGATCC	1994
QY	1791	TCTTTCAATCAGCCTTTCACAAATCCCATCCCATCTTAACCAAGTGGCCGCTTCTATC	1850
Db	1995	TCCTTCTTCACCTTTGTGATGACCCCATTTATCTTAACCAAGTCAGTCTTCTCTTTTC	2054
QY	1851	ATTCGCGAGTCCCTCTCGCGGTGAACCAATTAACCAACATGTCCTTCAATTTCCCTTT	1910
Db	2055	ATTCGCGAGTCCCTCTCGAGTTGGTGATCCACAGATAT---CTCCACAACTTCCCTTT	2111
QY	1911	CCTGAGAAATGGTGGCATTTCTTCTTCTTCTTCCCTCAGAACTCACAACAGTTGTC	1970
Db	2112	CCTGAGAAATGGTGGTATTTCTTCTTCTTCTTCTTCACTTGAATCTCAACAGTGTCC	2171
QY	1971	AGTTAACCCACCGGTCTCATGCGCGACTTTTAAGAACCCGAAACCGTGGCTAACTGGGA	2030
Db	2172	TGTCACTCTCCACAGTGTATCCCAACTCTTAAGAAACCCAAACCATTTGCTTAATGGGA	2231
QY	2031	ATCTATGCTTAAGCAATCATAGGCAAT---TGCTAAACATCAATGAGCGCTTTTAATTA	2087
Db	2232	ATCTTTTACCAACATCATATGTCCATGGCTCTTAACAGTCATATGATCTTTTGAACTA	2291
QY	2088	TCCTTTCTATGGGGTTTCTGCACTCTGTAGTCCGACACATGGCCACCAAGTTCAATCCT	2147
Db	2292	CCCGTTTATGGGGTGTCTGCACTGCAAGTCTCACTCATCATCGCCAGTTCATGCTCT	2351
QY	2148	GGCTACTATACCTGAATGTGAAGTCTGACCTTCCACTGTGATTTCTGGTCAATTTGAT	2207
Db	2352	GGCTACTATACCTGAATGTGAAGTCTGACCTTCCACTGTGATTTCTGGTCAATTTGAT	2411
QY	2208	AAGCTTTCAGAAAGTTTGCAACAACAGCAATTCCTGCGCTTAATGATGGCCAACTCTCC	2267
Db	2412	AAGCTTTCAGAAAGTTTGCAACAACAGCAATTCCTGCGCTTAATGATGGCCAACTCTCC	2477
QY	2268	TACCTTCAATCTTGTGAACCTGCGCTTCAGCAGATGTCTCCAAATCTGTGCTTCCA	2327
Db	2472	TACCTTCAATCTGCGGAACCTGCAACACAGCAAGATGTCTCCAAACACAGCAGCAATCCA	2531
QY	2338	AGAGATTGGTCAAAGCTCTGAATTTAAATTTGAGAAATAGCCAAAGTTAAACCTGGGAAG	2387
Db	2532	AGAGATTGGTCAAAGCTCCGAAATTTAAAGTTTGAAGAACCCAAAGTTAAAGCCATGGGAAG	2591
QY	2388	AGAGAGATACATGATGTGGGATGAGAGATCTTGAGCTTACACTGTGAAGAAATGGGAAGG	2447
Db	2592	GGAGAGATCCATGATGTGGGATGAGAGATCTTGAAGCTCAGGCTTGGAAATGTTAAGC	2651
QY	2448	TCGTGTTGACATTAACAACTAGGCAAAACCAATGGCATGTC	2490
Db	2652	TCATAGTTGAGATGAAGTATACATGAACCTGTATGTATCATGTC	2694
RESULT 6			
US-09-995-938A-3			
Sequence 3, Application US/09995938A			
Publication No. US20030150026A1			
GENERAL INFORMATION:			
APPLICANT: JOANNE CHORY			
APPLICANT: ZHIYONG WANG			
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID			
TITLE OF INVENTION: HORMONE ACTION IN PLANTS			
FILE REFERENCE: SALKINS_046A			
CURRENT APPLICATION NUMBER: US/09/995,938A			
CURRENT FILING DATE: 2001-11-27			

```

; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: ARABIDOPSIS THALIANA
; US-09-959-928A-3

```

Query Match	26.5%	Score 711.4;	DB 13;	Length 3000;
Best Local Similarity	71.1%;	Pred. No. 1.6e-161;		
Matches 1097;	Conservative 0;	Mismatches 376;	Indels 70;	Gaps 9

Qy	958	AGCAGAGAAAGGGATTTCTCGTGGACATAATTCACATCTCTCTTCTTCTT	1010
Db	1212	AGAAAAAAGGATATCTCTGTAGACATACTTCACATCA----TTCTTCTTCTT	1267
Qy	1018	CATCATCTAGTTCACACATCTTTACCCACCTATTTCAAAGCTCTCTCCGAAATTTT	1077
Db	1268	CAGCTGAATCCAAATACCCCATTTCCATTTTAACTGAGTTGATTTTGGAGATTTGA	1327
Qy	1078	GAGGGGTTGGTTGTGGTTTTCCGATGACTTGGATGGAGCTAACGTGCAATCAGACG	1137
Db	1328	AGGAAGAAG-----ATGAGCTCTAGGAGCAACGTGACGT--CAGC	1368
Qy	1138	TGCAGCTGCGGCGGCGCAGACGCGGAGAGGAGACCCTGTGAGAGAAAAGGAGAA	1197
Db	1369	TGCAGCTGCGAGACAGCATGGCGAGAGAGGAAACCTGCGTGGAGAGAGGAGAA	1428
Qy	1198	TAAATCGAGAGAGAAAGACGAGAGAGCTGTATGCTCGAAGATATACATCGGCTTAG	1257
Db	1429	CAATCGGAGAGAGAGCGGCGAGAAAGCTGTGTGCGGCAAGATTTATCTGCTCTTAG	1488
Qy	1258	AGCTCAAGGTGATTTATTTTGGCTTAACATGTGATTAATATGAAGTCCCTTAAGCT	1317
Db	1489	AGCTCAAGGTATCATCAATCTTCCAAAACATGTGACACAAATGAGTCTTAAAGCTCT	1548
Qy	1318	TTGTGTGAAGCTGGTGGGTGGTTGGTGAAGAAAGATGCTACTACTATCGCAAGTGAAGAC	1377
Db	1549	TTGTTCGTGAAGCTGGTGGGTGGTTGGTGAAGAAAGACGAACTACTATTCGCAAGTCAAT	1608
Qy	1378	TTTTCATCTTTTTCAGATCTGAGCTGTGTTATTTGATGTTTTTGAATCTGA	1437
Db	1609	CAACGCAATTTTAC---TTAGATCTGATGTTATCTGATGATTTAGCTGTGTAATCTGA	1664
Qy	1438	ATTGCTGATTTCAATGTGTGTTAAATGGGTTTGAAT---CTGGAATTTGAGGTTTTC	1494
Db	1665	AGATTGGATTGAAATTTGGTCAAAATTTGGGATTTTCTGGTATGAATTCAGGTTTTTA	1722
Qy	1495	TC---AAAGTGAATTTGAATCATCAGAAACTATGATGATGATTTCTCAAAAGTGA	1550
Db	1725	GCTGAGGAAGCTCAAGTTTTATTTCTTAAATTTGGAATTCGAGATCTCTTGGCGAAGAACTGAC	1784
Qy	1551	TTTATGGCTTTCTTCTTAATTTTGAAGTTATTTATGGTATGCTAAAGCTTAAATCTTTT	1610
Db	1785	TTTATGGCTTTCTTCTTAATTTTGAAGTTATTTATGGTATGCTAAAGCTTAAATCTTTT	1844
Qy	1611	ATGTATGATCTTGGTCCAAAGTCAATGATGATGTGTTCTTTTGGCTTACCTGTGATGAT	1670
Db	1845	ATTTA-----GATCTCTCTTAAAGCTTTGATTTGA	1874
Qy	1671	TGATCTTTGATGGTTATTTGTTTGGCTTTTGTGGAGATCAGGGATCAAGCTTTTAC	1730
Db	1875	TTGTGACATGTGATGTTTTTTGGTATTTGTTCTATATGAGACAAAGGACAAAGCTCTAC	1933
Qy	1731	TGCTGAGATAGTGGGACTTCAATCTCGAGTAACTCCATATTTCAACAAGCCGCTCTTCTTATC	1799
Db	1935	TGCTGACATAGCTGTGATCATCTTCGAGCAACTCTTACTCTTCCATTAACAAGTCC	1994
Qy	1791	TCCTTTCAATAGCTTTCAAAGTCCCATCTTCAACCAAGTTAGCCGCTCTTCTTATC	1855
Db	1995	TCCTTTCTTCACTTTTGAATAGCCCATCTTATCTTACCAAGTCACTCTTCTTCTTCTC	2054
Qy	1851	ATTCGAGTCTTCTCGCGGGAACAATTAACAATGTCCTTCAATCTTCTCTTT	1910


```

Db      2055 ATTCGAGAGTCTTCTTCGAGTTGGTGATATCCACAAATAT--CTCCAAATTTTCCCTTT 21111
QY      1911 CCTCAGAAATGGTGGCATTCTCTTCTTCTTCTTCCCTCAGAAATCTCAACAGTTGTC 19707
Db      2112 CCTCAGAGATGGTGGATATCTTCAATCGCTTCTCTCACTTAGAATCTCAACAGTGTCC 21717
QY      1971 AGTTAACCACCGGCTCTATCGCCGACTTTAAGAACCGGAAACCGTTGCTTAATGGGA 20336
Db      2172 TGTCACTCCACCAAGTGTCATCCCAACTTCTAGAAACCCCAACCAATGCTCACTTGGGA 22313
QY      2031 ATCTATCGCTAAGCAATCATGAGCCAT--TGTAAACAATCATGGGCTTTAATTA 20877
Db      2232 ATCTTTTACAAACATTCATGTCATGCTCTTAAACATCAATGAATCTTTTAATCTA 22911
QY      2088 TCCTTTCTATGCGGTTTCTGCACCTGTAGTCCGACACATCGCACCAAGTTTCATACCCT 21474
Db      2292 CCGGTTTATGCGGTGTCTGCACCTGCCAGTCTCTATCATATGCGCAGTTCCATGCTCC 23517
QY      2148 GGCTACTATACCTGAATGTGATGATGTGATCTGTCACTCTTCACTGTGTAATCTGGTCATTGGAT 22077
Db      2352 GGCTACTATACCTGAATGTGATGATGTGATCTGTCACTCTTCACTGTGTAATCTGGTCATTGGAT 24111
QY      2208 AAGCTTTCGAGAGTTTGCACAAACAGCATTTCTGTGCTGTATAGTGCCAACTCTCC 22677
Db      2412 AAGCTTTCAAAAGTTTGCACAAACAGCATTTCTGTGCTGTATAGTGCCAACTCTCC 24711
QY      2268 TACCTTCAATCTTTGTAAACCTGCGCTCAGCAGATGTCTCCAAATAGCTGACCTTGCA 23277
Db      2472 TACCTTCAATCTCGTGAACCTGCAACAGCAAGAAATGTCTCCAAACAGCAGCAATCA 25311
QY      2338 AGAGATTGGTCAAAGCTCTGAGATTTAAATTGGAATACCAAGTTAAACCTGGGANG 23877
Db      2532 AGAGATTGGTCAAAGCTCTGAGATTTAAATTGGAATACCAAGTTAAAGCATGGGANG 25911
QY      2388 AGAGAGATACATGATGTGGGATGTGAGAGATCTTAGGCTTACACTTGGAAATGGGAAGC 24477
Db      2592 GAGAGAGATCATGATGTGGCTATGAGAGATTTAAGGCTCAAGCTTGGAAATGGGAAGC 26511
QY      2448 TCGTGTTGACATTAACAATAAGCAAAACCAATGAGCATGTC 2490
Db      2652 TCATAGTTGAGATGAAGTATACATGAACCTGTATGTATGTC 2694

RESULT 7
US-09-995-938A-12
; Sequence 12, Application US/0995938A
; Publication No. US20030150026A1
; GENERAL INFORMATION:
; APPLICANT: JOANNE CHORY
; APPLICANT: ZHIYONG WANG
; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
; FILE REFERENCE: HORMONE ACTION IN PLANTS
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FASTSEQ for Windows Version 4.0
; LENGTH: 1213
; TYPE: DNA
; ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-12

```

Query Match	19.9%	Score 535.6	DB 13	Length 1213
Best Local Similarity	82.1%	Pred. No. 3,76-119		
Matches	642	Conservative	0	Mismatches 134; Indels 6; Gaps 2;

QY	1712	AGGATGCAAGCCTTACTGTGTGATGATAGCTGGGACCTTCATCTCGAATACTCCATT	1711
DB	409	AGGACACCAAGCTTACTGTGTGACATGGCTGGATCATCTTCTCGAACAACCTCTTAACT	468
QY	1772	CATCAGAACCAAGCCCTCTTTCATCAGAGCTTTCAAAGTCCATCCCATCTTATCCAG	1831

Db	469	CTTCCCATACCAAAAGTCCTCTTTCTTCACATTTTGATAGCCCATCTTATCTTAACCAAG	528
Qy	1832	TTAGCCGGCTCTTTCATCATTTCCCGAGTCCTTCGCGGGAACCAATAACAATGT	189
Db	529	TCAGTCCCTCCTCTTCTTCATTTCCCGAGTCCTTTCGAGTTGGATCCACACATAT-	586
Qy	1892	CCTGTACATTTTCCTCTTCTCGAATAAGTGGCATTCCTTCTCTTCCTTCCCTCA	1951
Db	587	-CTCCACATTTCTCCTTCTCGAAGATGGTATTTCTTATCGCTTCTCACATTA	645
Qy	1952	GAATCTCAAAAGTTTTCAGTTACCCACCGGTCTCATGGCCGACTTTAAAGACCGCA	2011
Db	646	GAATCTCAAAAGTGTCTCTGTCACTCCACAGATGATCCCCAATTCGTGAAGACCCA	705
Qy	2012	AACGGTGGCCTAATCGGAATCTATCCGTAAAGCATCATGGCCAT--TGCTAAACAT	2068
Db	706	AACCATTTGCCATCTTGGAAATCTTTTACCAAAACATCATATGTCATGAGCTCTAAACGT	765
Qy	2069	CAATGGAGCTTTTAAATATCTTCTATGCGGTTTCTGCACTGTAAGTCCGACATC	2128
Db	766	CAATGACTCTTTTAACTAACCTACCCGTTTATGCGGTCTGCACTGCGCATCTACTATC	825
Qy	2129	GCCACCAATTCATACCCCTGGCTACTATACCTGAATGTGATGACTGACTCTTCACTG	2188
Db	826	ATGCCAGTTCCAGTCTGGCTACTATACCTGAATGTGATGACTGACTCTTCACTG	885
Qy	2189	TTGATTTCTGGTCAATGGATTAAGCTTTCAGAAATTTGCAACAACAAGCATCTCTGCT	2248
Db	886	TTGATTTCTGGTCAATGGATTAAGCTTTCAGAAATTTGCAACAACAAGCATCTCTGCT	945
Qy	2249	CTATGGGCAACCTCTCCATCCTTGAATCTTGGAAACCTGGCCCTCAGAGATGTCTC	2308
Db	946	CTATGGGCAACCTCTCCATCCTTGAATCTTGGAAACCTGGCAACAAGAAATTTGCTC	1005
Qy	2309	CAATTACTGTCTCTTCCAGAGATTTGGTCAAAAGCTCTGAGTTTAAATTTGAGATAGCC	2368
Db	1006	CAACACAGCAGCATCCAGAGATTTGGTCAAAAGCTCTGAGTTTAAATTTGAGAACAGCC	1065
Qy	2369	AAGTTAAACCTTGGGAAGAGAGAGATATCATGATGTGGTATGAGAGATCTTGAGCTTA	2428
Db	1066	AAGTTAAGCATATGGGAAGGGAAGAGATCATGATGTGGCTATGAGAGATCTTAGAGCTCA	1125
Qy	2429	CACTTGGAATTGGAGAGCTCTGTGTGACATAACAACCTAAGGCAACCCAAATGGCATG	2488
Db	1126	CGCTTGGAAATGTAAAGCTCATATGTGAGATGAAAGTATACATGAACTGTATATGTCATG	1185
Qy	2489	TC 2490	
Db	1186	TC 1187	

```

RESULT 8
US-09-995-938A-11
; Sequence 11, Application US/09995938A
; Publication No. US2003015002641
; GENERAL INFORMATION:
; APPLICANT: JOANNE CHORY
; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
; TITLE OF INVENTION: HORMONE ACTION IN PLANTS
; FILE REFERENCE: SALKINS. 046A
; CURRENT APPLICATION NUMBER: US/09/995,938A
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1213
; TYPE: DNA
; ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-11

```

Query Match 19.9%; Score 534; DB 13; Length 1213;

Best Local Similarity 82.0%; Pred. No. 9.1e-119;
Matches 641; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

```
QY 1712 AGGATGCAAGCCCTTACCTGGTGAATAGTGGACTTCACTCGAGTAATCCCATTT 1771
DB 409 AGGATGCAAGCCCTTACCTGGTGAATAGTGGACTTCACTCGAGTAATCCCATTT 468
QY 1772 CATACAGAAAGCAAGCCCTTTCATCAGCCTTCAAAAGTCCCATCCCATTAACAG 1831
DB 469 CTTCCCATTAACCAAGTCTCTTCTTCACCTTTGATAGCCCATCTTATCTTAACCAAG 528
QY 1832 TTAGCCGCTCTTCTTCATCATTCGAGTCTCTCGGGTGAACCAATATAACATGT 1891
DB 529 TCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 586
QY 1892 CCTCTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1951
DB 587 -CTCACAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 645
QY 1952 GAATCTCAAAAGTGTCTCACTTACCCCAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 2011
DB 646 GAATCTCAAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 705
QY 2012 AACGCTGCTTACCTGGAAATCTATCGTAAAGCAATCAATGAGCAT--TGCTAAACAT 2068
DB 706 AACCAATGCTTACCTGGAAATCTTACCAACCAATCAATGAGCATGAGCTGCTAAACAT 765
QY 2069 CAATGAGCTCTTTTAAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2128
DB 766 CAATGAGCTCTTTTAAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 825
QY 2129 GCCACAGCTTCAATACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2188
DB 826 ATCCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 885
QY 2189 TTGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2248
DB 886 TTGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 945
QY 2249 CTATGCTGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2308
DB 946 CTATGCTGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1005
QY 2309 CAATATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2368
DB 1006 CAATATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1065
QY 2369 AAGTTAAACCTGGAAGAGAGAGATCAATGATGATGATGATGATGATGATGATGATG 2428
DB 1066 AAGTTAAACCTGGAAGAGAGAGATCAATGATGATGATGATGATGATGATGATGATG 1125
QY 2429 CACTTGAAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 2488
DB 1126 CACTTGAAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1185
QY 2489 TC 2490
DB 1186 TC 1187
```

RESULT 9
US-09-938-842A-385
Sequence 385, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: S01P1300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A

CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 385
LENGTH: 831
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-385

Query Match 5.5%; Score 149; DB 10; Length 831;
Best Local Similarity 76.8%; Pred. No. 1.7e-25;
Matches 182; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```
QY 1134 CAGCTGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1193
DB 5 CGCATCAGAGAGAGAGATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 64
QY 1194 AGAATATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1253
DB 65 AGAATATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 124
QY 1254 TTAGAGCTCAAGCTGATTAATTTGCTTAAACATGATGATTAATGATGATGATGATG 1313
DB 125 TTAGATCTCAAGAGCAATTAATTTGCTTAAACATGATGATTAATGATGATGATGATG 184
QY 1314 CTCTTGTGTTGAAGCTGTTGGTTGTTGAAGAGATGATGATGATGATGATGATGATGAT 1370
DB 185 CTCTTGTGTTGAAGCTGTTGGTTGTTGAAGAGATGATGATGATGATGATGATGATGAT 241
```

RESULT 10
US-09-878-574-5002
Sequence 5002, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21 (15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 5002
LENGTH: 413
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: LIB3028-009-Q1-B1-D2
US-09-878-574-5002

Query Match 4.5%; Score 121.4; DB 10; Length 413;
Best Local Similarity 69.8%; Pred. No. 5.3e-19;
Matches 164; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
QY 1136 GCTGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1195
DB 92 GCCGAAACATGATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 151
QY 1196 AATATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1255
DB 152 AACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 211
QY 1256 AGAGCTCAAGGATGATTAATTTGCTTAAACATGATGATTAATGATGATGATGATGAT 1315
DB 212 CAGGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 271
```

2430 61.001.00M1.1CMM.101001.100A10001.10/M1.10M0M1.10M0001.11C1CM 2437

1438 ATTCGTGATTCAATTGTGGTTAAATGGGTTTGAATCTGAGAAATTGAGGGTTTCTCA 1497

[illegible][illegible]

Qy	1696	CTTTTGT	1703
Db	971860	TTTTTTTTT	971867

```

RESULT 15
US-10-311-455-330
; Sequence 330. Application US/10311455
; Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
TITLE OF INVENTION: Cytosine methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 330
; LENGTH: 9642
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-330

```

Query Match	Best Local Similarity	2.3%;	Score 62.6;	DB 13;	Length 9642;
Matches 161;	Conservative 0;	Mismatches 164;	Indels 0;	Gaps 0;	
QY	1386	TTTTTCCAGATCTGAGCTGTTTATATGATGTTTTGATGTTTGAATCTGAATTCGTG	1445		
DB	3491	TTTTTTTTTATTAATAGTTTTTTAAAGAAATAGCTTATATTTGTGATTTTATATTTTGA	3550		
QY	1446	ATTTCAAATGTGCTTAAATGGGCTTTGAATCTGGAATTTGAGGCTTTTCCTCAAGTGAAT	1505		
DB	3551	TTTTTATTTTTTATTAATTTTTTATTTGGCTTTTATTTATATATTTTGTGCAAAATTCGT	3610		
QY	1506	TTGAATCATCAGAAACTATGATGATCGATCTGATTTCTCAAGTGAATTTATGGCTTTCTT	1565		
DB	3611	TAAATTATTTAATGATGATATATTTTGTATATTTTCGTAGATTTAAATTTTTTAAATATATTT	3670		
QY	1566	TCTAATTTTAGAGTATTAATTTGGTATGCTAAAGCTTAACTTTATATGATGATCTTGG	1625		
DB	3671	TTAATATTAGGGGTTTTTTTTTAATATTTTATATATTTTTCGTAAATTTTTTTTTTTAGAACT	3730		
QY	1626	TCCAAAGTCATTCATGTGCTTTCTTTGGCTTAACCTGATGATGATGATGATTTGGT	1685		
DB	3731	TTTTTATTTTTTTGGCTTTTAAAGGTAGTGAATATTTTGAATATTTTATTTTGAATTT	3790		
QY	1686	TATGTTTTGCTTTTGGTGGAGTAT	1710		
DB	3791	TTTTTTATTTTTTTTTTGGAAAGT	3815		

Search completed: December 20, 2003, 03:48:24
Job time : 813 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 20:38:22 ; Search time 9603 Seconds

(without alignments)
11446.857 Million cell updates/sec

Title: US-09-995-938a-2

Sequence: 1 acttactgttcgcctcaattc.....attatttgtagataaaaaa 2687

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pac.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_srb.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: gb_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_ov.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_srb.*

28: em_un.*

29: em_vl.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_ocher.*

33: em_hg_mus.*

34: em_hg_pln.*

35: em_hg_rod.*

36: em_hg_man.*

37: em_hg_vte.*

38: em_ey.*

39: em_hgo_hum.*

40: em_hgo_mus.*

41: em_hgo_ocher.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2685.4	99.9	2687	8	AF494338	AF494338 Arabidops
2	2685.4	99.9	97263	8	AC013258	AC013258 Arabidops
3	1088.8	40.5	1476	8	AY065049	AY065049 Arabidops
4	1068.2	39.8	1501	8	AY087257	AY087257 Arabidops
5	742.8	27.6	1011	8	AY093747	AY093747 Arabidops
6	711.4	26.5	120977	8	AC025808	AC025808 Genomic s
7	534	19.9	1340	8	AY065041	AY065041 Arabidops
8	532.4	19.8	1359	8	AY086340	AY086340 Arabidops
9	530.8	19.8	1257	8	AF134217	AF134217 Arabidops
10	501.4	18.7	945	8	AF372937	AF372937 Arabidops
11	326.4	12.1	411	8	AY074829	AY074829 Arabidops
12	194.8	7.2	1494	8	AF395901	AF395901 Lycopersi
13	156	5.8	19854	8	ATAP22	299708 Arabidops
14	156	5.8	198780	8	ATCHR186	AL161590 Arabidops
15	152.6	5.7	84196	8	ATT3A5	AL133979 Arabidops
16	152.6	5.7	89469	8	ATF1883	AL049862 Arabidops
17	149	5.5	831	6	AX505690	AX505690 Sequence
18	149	5.5	831	6	BT006310	BT006310 Arabidops
19	149	5.5	1273	8	BT002452	BT002452 Arabidops
20	142	5.3	1889	8	AY050394	AY050394 Arabidops
21	141.4	5.3	957	8	AY097357	AY097357 Arabidops
22	118	4.4	90425	8	P9K20	AY090331 Arabidops
23	117	4.4	978	8	AY090331	AY090331 Arabidops
24	117	4.4	1515	8	AY050430	AY050430 Arabidops
25	113.2	4.2	11911	8	ATRI3C5	AL021711 Arabidops
26	113.2	4.2	199075	8	ATCHR149	AL161549 Arabidops
27	110.8	4.1	855	6	AX651445	AX651445 Sequence
28	110.8	4.1	1217	8	AY088379	AY088379 Arabidops
29	110.8	4.1	1258	8	AK118850	AK118850 Arabidops
30	110.6	4.1	14480	8	AP004276	AP004276 Oryza sat
31	108.2	4.0	897	6	AX653881	AX653881 Sequence
32	104	3.9	158815	8	AP003105	AP003105 Oryza sat
33	96	3.6	166126	2	AP003686	AP003686 Oryza sat
34	93.4	3.5	1116	6	AX653341	AX653341 Sequence
35	93.2	3.5	192110	2	AP004070	AP004070 Oryza sat
36	92.4	3.4	1374	8	AY200612	AY200612 Arabidops
37	85	3.2	176734	2	AP003682	AP003682 Oryza sat
38	77.6	2.9	2070	8	BT006482	BT006482 Arabidops
39	77.6	2.9	2272	8	AK117140	AK117140 Arabidops
40	77.6	2.9	62052	8	AB020744	AB020744 Arabidops
41	71.4	2.7	349980	6	AX344560	AX344560 Sequence
42	61.45	2.6	6145	6	AX345874	AX345874 Sequence
43	69.6	2.6	5499	6	AX458455	AX458455 Sequence
44	67	2.5	109135	8	AP003408	AP003408 Oryza sat
45	67	2.5	155955	8	AP003288	AP003288 Oryza sat

ALIGNMENTS

RESULT 1

AF494338 2687 bp DNA linear PLN 20-MAY-2002

LOCUS Arabidopsis thaliana BZRI protein (BZRI) gene, complete cds.

DEFINITION AF494338

ACCESSION AF494338

VERSION AF494338.1 GI:20270970

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE 1 (bases 1 to 2687)

Wang,Z.Y., Nakano,T., Gendron,J., He,J., Chen,M., Vafeados,D., Yang,Y., Fujikura,S., Yoshida,S., Asemit,T. and Chory,J.

TITLE Nuclear-localized BZR1 mediates brassinosteroid-induced growth and feedback suppression of brassinosteroid biosynthesis
JOURNAL Dev. Cell 2 (4), 505-513 (2002)
MEDLINE 21966115
PubMed 11970900
REFERENCE 2 (bases 1 to 2687)
AUTHORS Mang, Z.-Y. and Chory, J.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2002) Plant Biology, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Rd., La Jolla, CA 92037, USA

FEATURES
source location/qualifiers
1..2687
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/map="between ngal11 and AcatPase"
/note="ecotype: Columbia"
1..2687
/gene="BZR1"
1..880
/gene="BZR1"
join(881..1369,1714..2687)
/gene="BZR1"
/product="BZR1 protein"
join(1103..1369,1714..2457)
/gene="BZR1"
/codon_start=1
/product="BZR1 protein"
/protein_id="AA18490.1"
/db_xref="GI:20270971"
/translation="MTSDGATSTSAIAAAAAAAAAARRKPSWRENNRRRRRAVA
AKIYTGAAQGVNLPKHCNNEVYLKALCYAGVNEDEGTYTKGCKPLGEIAGTS
SRVPSYSONOSPLISAFOPIPSYOVSPSSPSPSRGRPNMMSSTPEPIRNG
IPSELPRLRISNSCPVTPVSPSPSKPKPLPNMSEIAKOSMAIAKSMASFNTPFA
VSAASPTHRHOFHPATIPATIPEDSDSDSTDSGHMISFOKAQCPFSASAVPISPTF
NLVVPAPQMSPTNPAFOEIGSSSEFFKENSQVPMWEGRIHDVGMDELDTLGNKRA
RG"
2147
/gene="BZR1"
/note="compared to BZR1-1D mutant"
/replaces="c"
BASE COUNT 781 a 554 c 494 g 858 t

variation
Query Match 99.9%; Score 2685.4; DB 8; Length 2687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2686; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTTTAGTTTGCCTTAATTCATCGAACCTCTGATTCATTCCAATGTTTCCCAACTGCG 60
Db 1 ACTTAGTTTGGCTTAATTCATCGAACCTCTGATTCATTCCAATGTTTCCCAACTGCG 60
QY 61 GTTGATGTTGGGTTCTCTGCTTTTAAGACATTAAGTACATGCTCTGCACTTTTAC 120
Db 61 GTTGATGTTGGGTTCTCTGCTTTTAAGACATTAAGTACATGCTCTGCACTTTTAC 120
QY 121 AAAATAAAGCTCATATTCGAAATAATATACATGATCATATATATGCGCGGAAC 180
Db 121 AAAATAAAGCTCATATTCGAAATAATATACATGATCATATATATGCGCGGAAC 180
QY 181 TTGTTATGGGACAACTGTAACCCCTTTTCTTTATGTTCAATGAATATACAGT 240
Db 181 TTGTTATGGGACAACTGTAACCCCTTTTCTTTATGTTCAATGAATATACAGT 240
QY 241 TTGTTATGGGACAACTGTAACCCCTTTTCTTTATGTTCAATGAATATACAGT 300
Db 241 TTGTTATGGGACAACTGTAACCCCTTTTCTTTATGTTCAATGAATATACAGT 300
QY 301 GATACATAAGCTTAAATCAACATGTAACCAAACTAAATACCTTATTAAGACATGTAAT 360
Db 301 GATACATAAGCTTAAATCAACATGTAACCAAACTAAATACCTTATTAAGACATGTAAT 360

QY 361 GGTATTCACCAATCTTTATATCATTTGTAAAGTACGAAGAGTAAAAAAGACAGAGC 420
Db 361 GGTATTCACCAATCTTTATATCATTTGTAAAGTACGAAGAGTAAAAAAGACAGAGC 420
QY 421 CAGGTACATACAACTAATCAGGACAAAGTACAAATGTTTCTTAACTGAGATTG 480
Db 421 CAGGTACATACAACTAATCAGGACAAAGTACAAATGTTTCTTAACTGAGATTG 480
QY 481 TATGCAAGAAAAAGTATATTTTAAATGAATATATCATTTATGATGTTATACACGG 540
Db 481 TATGCAAGAAAAAGTATATTTTAAATGAATATATCATTTATGATGTTATACACGG 540
QY 541 CTTCACGTATATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
Db 541 CTTCACGTATATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
QY 601 AGTAAGGAGCCACACGTAAGAAACGGTCCACGTGTCTCTCTTCTTCTCTGTA 660
Db 601 AGTAAGGAGCCACACGTAAGAAACGGTCCACGTGTCTCTCTTCTTCTCTGTA 660
QY 661 TTATTTGTCATTAATCAATTTTATACATCTCACAGAAAAAAGATTAGAACATAAC 720
Db 661 TTATTTGTCATTAATCAATTTTATACATCTCACAGAAAAAAGATTAGAACATAAC 720
QY 721 ACAGTTACTAAGGATATATCTCTGACCTTAACATACACCTTTATATTCACCTCA 780
Db 721 ACAGTTACTAAGGATATATCTCTGACCTTAACATACACCTTTATATTCACCTCA 780
QY 781 CGTAACTCAACCTTCCAAAAACATGATTTTACACGTGAGACATGATGATACCAAGACA 840
Db 781 CGTAACTCAACCTTCCAAAAACATGATTTTACACGTGAGACATGATGATACCAAGACA 840
QY 841 TGATTTTAATATGAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 TGATTTTAATATGAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 TGAATAAATAATATATTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 TGAATAAATAATATATTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 GAAGGAAAGGATATCTCTGAGACATATCTCTGAGACATATCTCTCTCTCTCTCTCT 1020
Db 961 GAAGGAAAGGATATCTCTGAGACATATCTCTGAGACATATCTCTCTCTCTCTCTCT 1020
QY 1021 CAGTCTAGTTCACACATCTTTTCAACCACTTAAAGCTCTCTCCGAAATTTCCAG 1080
Db 1021 CAGTCTAGTTCACACATCTTTTCAACCACTTAAAGCTCTCTCCGAAATTTCCAG 1080
QY 1081 GGGTTGGTTGGTTTCCCGATGACTTGGATGAGACTTACGTCGATCAGACAGCTGC 1140
Db 1081 GGGTTGGTTGGTTTCCCGATGACTTGGATGAGACTTACGTCGATCAGACAGCTGC 1140
QY 1141 AGCTGCGCGCGCGGACACACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 AGCTGCGCGCGCGGACACACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 TCGAG 1260
Db 1201 TCGAG 1260
QY 1261 TCAAGGTGATTAATATTTGCTTAACATTTGATTAATTAATGAAGCTCTTAAAGCTTT 1320
Db 1261 TCAAGGTGATTAATATTTGCTTAACATTTGATTAATTAATGAAGCTCTTAAAGCTTT 1320
QY 1321 TGTGAAGCTGGTGGTGGTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 TGTGAAGCTGGTGGTGGTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 CTCCATTTTTCAGATCTGAGCTGTTTATATATGTTTGTGAATGTTGAATCTGAAT 1440
Db 1381 CTCCATTTTTCAGATCTGAGCTGTTTATATATGTTTGTGAATGTTGAATCTGAAT 1440

1441 CGTTGATTTCAATGCTTAAATGGGTTGAATCTGAGAAATTTGAGGTTTCTCAAG 1500
 1441 CGTTGATTTCAATGCTTAAATGGGTTGAATCTGAGAAATTTGAGGTTTCTCAAG 1500
 1501 TGAATTTGATCATCAGAAACTATGATGATGATGATGATGATGATGATGATGATGAT 1560
 1501 TGAATTTGATCATCAGAAACTATGATGATGATGATGATGATGATGATGATGATGAT 1560
 1561 TTTCTTTCTAATTTTGAAGTATTAATGATGATGATGATGATGATGATGATGATGATGAT 1620
 1561 TTTCTTTCTAATTTTGAAGTATTAATGATGATGATGATGATGATGATGATGATGATGAT 1620
 1621 CTGATGTCAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 1621 CTGATGTCAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 1681 TTGGTATTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1740
 1681 TTGGTATTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1740
 1741 GCTGGGACTTCATCTGAGTAACTTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1800
 1741 GCTGGGACTTCATCTGAGTAACTTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1800
 1801 GCTTTTCAAGTCCCATTCCTTACCAAGTACCGCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 1801 GCTTTTCAAGTCCCATTCCTTACCAAGTACCGCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 1861 CCTTTCGGGGTGAACCAATATACCAATATACCAATATACCAATATACCAATATACCAATAT 1920
 1861 CCTTTCGGGGTGAACCAATATACCAATATACCAATATACCAATATACCAATATACCAATAT 1920
 1921 GGTGGCAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
 1921 GGTGGCAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
 1981 CCGGTCTCATGCGGACTTCTTCAAGAACCGGAAACCGGAAACCGGAAACCGGAAACCGGAA 2040
 1981 CCGGTCTCATGCGGACTTCTTCAAGAACCGGAAACCGGAAACCGGAAACCGGAAACCGGAA 2040
 2041 AAGCAATCCATGCGGACTTCTTCAAGAACCGGAAACCGGAAACCGGAAACCGGAAACCGGAA 2100
 2041 AAGCAATCCATGCGGACTTCTTCAAGAACCGGAAACCGGAAACCGGAAACCGGAAACCGGAA 2100
 2101 GTTTCGCACTGCTAGTCCGACATGCGACATGCGACATGCGACATGCGACATGCGACATGCGAC 2160
 2101 GTTTCGCACTGCTAGTCCGACATGCGACATGCGACATGCGACATGCGACATGCGACATGCGAC 2160
 2161 GAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
 2161 GAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
 2221 TTTTGCACAACAACGACATCTCTGCTCTTATGAGGCAACCTCTCTCAACCTCTCAACCTCT 2280
 2221 TTTTGCACAACAACGACATCTCTGCTCTTATGAGGCAACCTCTCTCAACCTCTCAACCTCT 2280
 2281 GTGAACCTGCGGCTCAGAGATGCTCTCAACCTCTCTCAACCTCTCTCAACCTCTCTCAACCT 2340
 2281 GTGAACCTGCGGCTCAGAGATGCTCTCAACCTCTCTCAACCTCTCTCAACCTCTCTCAACCT 2340
 2341 AGCTCTGAGTTTAAATTTGAGAAATGCGAAGTTAAACCTCTGGAAGGAGAGAGATACAT 2400
 2341 AGCTCTGAGTTTAAATTTGAGAAATGCGAAGTTAAACCTCTGGAAGGAGAGAGATACAT 2400
 2401 GATGCGGATGAGAGATTTTGAAGCTTCACTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAT 2460
 2401 GATGCGGATGAGAGATTTTGAAGCTTCACTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAT 2460
 2461 AAACAAGTGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2520
 2461 AAACAAGTGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2520
 2521 TATTTTCTTCTTCTGTCAGGATTTTGGATCTTATGAAATCTCAATATGTTCTTCACTTAT 2580

Db 2521 TATTTTCTTCTTCTGTCAGGATTTTGGATCTTATGAAATCTCAATATGTTCTTCACTTAT 2580
 Oy 2581 TATTCGAATATGCTGCTCCCAAGCCTTCTCATGAGAACATTTGAGGTTAGAGGTTAT 2640
 Db 2581 TATTCGAATATGCTGCTCCCAAGCCTTCTCATGAGAACATTTGAGGTTAGAGGTTAT 2640
 Oy 2641 TCAATTCATGATTTGTTTCAAGACATTTATTTGATTAATAAAAA 2687
 Db 2641 TCAATTCATGATTTGTTTCAAGACATTTATTTGATTAATAAAAA 2687

RESULT 2

AC013258/ 97263 bp DNA linear PLN 19-JAN-2001
 LOCUS Arabidopsis thaliana chromosome 1 BAC F9E10 genomic sequence.

AC013258 complete sequence.

AC013258.5 GI:12323880

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BAC clone F9E10 is from Arabidopsis thaliana chromosome 1

The orientation of the sequence is from Sp6 to T7 end of the BAC clone.

Genes were identified by a combination of several methods: Gene

prediction programs including GenScan+ (Chris Burge,

http://CCR-081.mtc.edu/GENSCAN.html), GenemarkMM (Mark Borodovsky,

http://genemark.biology.gatech.edu/Genemark/), Glimmer3 (a variant

of Glimmer, see Mikaela Pertea,

http://www.tigr.org/softlab/glimmer3.htm, and

GeneSplicer (Mikaela Pertea and Steven Salzberg, contact

mpertea@tigr.org), searches of the complete sequence against a

peptide database and the plant EST database at TIGR

(http://www.tigr.org/tcb/cgi.shtml). Annotated genes are named to

indicate the level of evidence for their annotation. Genes with

similarity to other proteins are named after the database hits.

Genes without significant peptide similarity but with EST

similarity are named as unknown proteins. Genes without protein

or EST similarity, that are predicted by more than two gene

prediction programs over most of their length are annotated as

hypothetical proteins. Genes encoding tRNAs are predicted by

tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).

Simple repeats are identified by RepeatMasker (Arian Smit,

http://ftp.genome.washington.edu/RM/RepeatMasker.html).

Location/Qualifiers

1. 97263

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/culliv="Columbia"

/db_xref="taxon:3702"

```

/chromosome="1"
/clonase="P9E10"
14..1286
/gene="P9E10.1"
<14..1286
/gene="P9E10.1"
14..1201
/gene="P9E10.1"
/codon_start=1
/product="unknown protein; 14-1201"
/translaction="MENCKPFGDTEGLOVTLFVHGVNRVATLDTSGKLTJFTE
NRTYGSVPTSRPLVFLKQRLFLFETGAGSLDRSKIRSECEGLNHSIARIVF
DESERKAYGFTSEGEIIVHLLGDIIMKCKVRSKQVMEPALQIKYLLIVN
QEKVFNVNSTOHWYRTGPRLLPPALIEDISTPLSHRESTKTDHOKLEVTDLIA
SDREKLWGLGDGVATYKSLPIKAEFNTLMSSPVFFILFLFGAMHFFSKKE
SLTAGPDPPFSTTMSSTTTAQNSSAFSESTRNDHMLRRRYVSPSRYPGAA
TGAYSVGSNDSSRAPVETTYRTTAQEMKRGSGLDGSGFGKRBSLFGNNALD
DES"
1456..3294
/gene="P9E10.2"
/note="similar to cytochrome P450 GB:AA17732
[Catharanthus roseus]; contains Pfam profile: PF00067
Cytochrome P450"
join(<1456..1944,2099..2343,2436..2799,2875..3294)
/gene="P9E10.2"
join(1456..1944,2099..2343,2436..2799,2875..3294)
/gene="P9E10.2"
/codon_start=1
/product="putative cytochrome P450; 1456-3294"
/protein_id="AAG51924.1"
/db_xref="GI:12323898"
/translaction="MAVFFILVLFVFFVFRFVYINWVPMRIQSHFKQSYTGPSYR
IFSGNSGEVSRLTAAKSPKIPSGRNPHEFVRVAPHYEMSRVAGKTELYFGSKPV
VATSDPRLIREALTGGSPDRIGHNPISLKLAAQIPGLRGDQMAFHRIRAKQARTME
KLKRVPMQVSTWMLMEKMDMRNGCEIELEHKKEMNLSEMLSRAPGNSYBEG
KGIPELOERMRILFYLRKSVYITPGRRFPSTKNEIMRIEKOIVSILKLENNKTA
VEKSTLLQAFMSPYTNQGOEKKIGIEVTECKTFPFAAETTMNTFVLVLLAM
NOEMQNIAREVIVCIQGTGLPTDLIDLKTLSMIINETLLVPPAMTLMNDTLKRA
KLGDIDIPAGTQILSVVAMHDKETMGDAEFPNRRDEPKQSALLVPGGLPRT
CVGONLAVNEAKTVALTIKYSFRLSPSYAANAPVTLQPNQAHLLFTTIS"
4711..6345
/gene="P9E10.3"
join(<4711..4845,5272..6345)
/gene="P9E10.3"
join(4711..4845,5272..6345)
/gene="P9E10.3"
/codon_start=1
/product="unknown protein; 4711-6345"
/protein_id="AAG51930.1"
/db_xref="GI:12323904"
/translaction="MAVRKEKQVPPREGIAIYAVLVGIFGVCITILLPNDPFRSS
KVASGSGSPERVKMFKAEPALISBKNELRQVSDLTREKVLAEOKVYKPGCVT
TGLQNTPTVAPDESANRLAKLEKAYAKKEITVLLANNVPMLEVOJASIKRQIGI
NVLVPLDDLESFCKSNVAYYKRDPSDAIDVAGSRSSDSGLKRVLEPFIQIG
YGVLLSDVDIVFLQNPGLYRSDVESDHDNNTYAGFNDVDDPMTSRFVQIG
NRIWFGSGFYLRLPTLSIELDRVDTLISKSGMDQAVFQHLFYPSHPTGLIYA
SKRVWDVYEFMNSRVLFTVRKDEEMKKLKPIYIHMYHSDLEEMQAAYEVYVNGKQ
DALRFRDGS"
7482..9019
/gene="P9E10.4"
join(<7482..7655,7907..9019)
/gene="P9E10.4"
join(7482..7655,7907..9019)
/gene="P9E10.4"
/codon_start=1
/product="unknown protein; 7482-9019"
/protein_id="AAG51917.1"
/db_xref="GI:12323891"
/translaction="MAGRRDRIOOLRGSRIFAIFVGLIGCVGLPDPNGFSSGSS
LIANEBSKSTSDGLASCSSESRVKMKPSIISYVNAALRQVVEITKVLAE
GETENARKOVLVLSSEITKAGPGLVAKSLTNPFTVDPDESVPRLAKLEKAVANKEII
```

```

VLANSNVKPMLEIOIASVKRYGIONVYLIALDSDMESFCEKEVYFKRDPDKAVDM
VKSQGNANAVSGIKRVRVREPIQIGSYVILSDVDIVFLQNPSSHLRSDVSMGCH
DNNTAYGFNDVDESPSGMAYAHYMRITWVSGFYLRTPIPSIDILDRVADTLKRS
EAMDQAVNEOLFYPSPHGYTGLHASKVMYDEPMNSKVLFTVRKONELKLEKVI
VHLNTHPDKLEMRHAIVEFYVNGKODALDSPDGS"
9598..12387
/gene="P9E10.5"
join(<9598..10007,10087..10712,10801..11195,11275..11389,
11467..11516,11613..11678,11776..11853,11944..12120,
12221..12387)
/gene="P9E10.5"
join(9598..10007,10087..10712,10801..11195,11275..11389,
11467..11516,11613..11678,11776..11853,11944..12120,
12221..12387)
/gene="P9E10.5"
/codon_start=1
/product="unknown protein; 9598-12259"
/protein_id="AAG51921.1"
/db_xref="GI:12323895"
/translaction="MOTLPSSSTVLLGSNSAPVLRSFGGDVDVIDFGVFGGPPRR
SKVTSNEVYTRHSFSESALRRRDVIVDQDLFQDEKPVGEDTSSVRRRTDPPDD
IFRVNESSILPGSRILSPAHKREBSGTSSPQSLPAKATEIPTFNLAATSLNKV
ETVSSPLSRSSKADVVSTAKSYGDDCDPPQVPTVGROQHFISLYKMPKGVAV
TMSGRILSSMKAEETTPVPLSDVKTSTVEVLEGNREBDSGSLGIDVKTSLKR
FGVQYKEKETETDLKSEQAPFGVGRKARANKVPLDSVBSQKPSGVSKAHEATTVP
HSIFHEBERQDEKIVSEREVKSKKANKRSTFEDSTTKRSQGTSLSSPIPD
KSSFASSSAAPEVGDGVKGVSDVKIPSKASVAGAGESAGOSRMRKAEKPTDI
IHDGNAKETVNIIPQOKKSTPDIIPAMRDQKOSTOKSDRESMYKAPEDTVOE
EROESTHTTISEDIDEPFVNFVDEDITODKNKEANKDAEINIDAKIRKMSG
KSGNIRSLSTLOYTIMSGMGPVPLMDMIEGNVRSYQALAILHDLKQOKKAS
ANQKYMAREVLELOEANDHPTLGPV"
complement(14275..14526)
/note="D11D1 Non-autonomous DNA transposon [a
consensus]."
14940..15720
/gene="P9E10.6"
/note="similar to putative DNA-3-methyladenine glycoylase
I (TMG1) GB:PO5100 [Bacterichia coli]"
join(<14940..15100,15179..15320,15400..15720)
/gene="P9E10.6"
join(14940..15100,15179..15320,15400..15720)
/gene="P9E10.6"
/codon_start=1
/product="putative DNA-3-methyladenine glycoylase I;
14940-15720"
/protein_id="AAG51925.1"
/db_xref="GI:12323899"
/translaction="WVIDOTDPIYVLPHDEEWVPRDDKLPILLVFSQALAEFSP
SILRRDPRKLFERPDPSAIAQFTKMLSRVNGCLILSEOKLAITVENKASVUKV
KOEFGSPSNYCMRFVFNHKLPLNNGYRGQVQVPKSKAEYISKDMQGRFCVGPVWY
SFLQASGIVNHLTACFRYOECNVETERTSHETETKLDLHSPLV"
complement(16304..17716)
/gene="P9E10.7"
complement(join(<16304..17038,17392..17716))
/gene="P9E10.7"
complement(join(16304..17038,17392..17716))
/gene="P9E10.7"
/codon_start=1
/product="unknown protein; 16304-17658"
/protein_id="AAG51929.1"
/db_xref="GI:12323903"
/translaction="MTSDGATSTSAAAAAAARRKPSWRENNRRRRRAVA
AKIYGLAAGQDYNLPKHCDDNNEVILKALCEVAGVVEEDGTYRKPLPEIAGTSRV
TPYSQONSPLSAFOSPISYOVSPSSS6PSPSPGNMNSSTPPEFLNGGSPS
SLPSLRISNSCPMPVPSVPSKSNKPIPMNYSIKOSMAIKOSMAISNPPYANSA
PASPTRHQFTPTATIEPCDSSDSTVPSGHHISIQKQAOOPFASASVPTSPYRLV
KPAQOWSPNTRAFQELIGQSSSEFPENQVKBWGERJHDVEMDELETLGNGKRG"
complement(18764..18835)
Query Match 99.9%; Score 2685.4; DB 8; Length 97263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2686; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```


Qy	1	ACTTTAGTTTGGCTTAATTCATGAAACCCCTGTGATTCATTCGAAATGTTTCCCAACCTCGC	60
Db	18760	ACTTTAGTTTGGCTTAATTCATGAAACCCCTGTGATTCATTCGAAATGTTTCCCAACCTCGC	18701
Qy	61	GTTGATGGTTCGGGTCCTCTGCTTTTAAACAATTAGTATCATGTCCTGTCGACTTTCTAC	120
Db	18700	GTTGATGGTTCGGGTCCTCTGCTTTTAAACAATTAGTATCATGTCCTGTCGACTTTCTAC	18641
Qy	121	AAATTAACGTCATAATCCAAAAATATATACATGATCATATATATGCGCGGAAC	180
Db	18640	AAATTAACGTCATAATCCAAAAATATATACATGATCATATATATGCGCGCGGAAC	18581
Qy	181	TGTGTTATGGGACAACCTCGTAAACCCCTTTTCTCTTTATGTTAAAGAACTATACAGT	240
Db	18580	TGTGTTATGGGACAACCTCGTAAACCCCTTTTCTCTTTATGTTAAAGAACTATACAGT	18521
Qy	241	TTTGGTTATGAATACATAAATTAATGATGAGCCCGCATTAATCAAAATTTGGATATTA	300
Db	18520	TTTGGTTATGAATACATAAATTAATGATGAGCCCGCATTAATCAAAATTTGGATATTA	18461
Qy	301	GATACTAAAGCTTAAATCAACATGTAACCAACTAAATCTTTATAGACATAGTAAT	360
Db	18460	GATACTAAAGCTTAAATCAACATGTAACCAACTAAATCTTTATAGACATAGTAAT	18401
Qy	361	GGTATTCACCAATCTTTATATCATTTGTAGGTACGAAGAGGTAAAAAGAGAGAGC	420
Db	18400	GGTATTCACCAATCTTTATATCATTTGTAGGTACGAAGAGGTAAAAAGAGAGAGC	18341
Qy	421	CAGTGACATACAACTAATACGAGCAAAAGTACAAATGGTCTTAAGTGAGATTGG	480
Db	18340	CAGTGACATACAACTAATACGAGCAAAAGTACAAATGGTCTTAAGTGAGATTGG	18281
Qy	481	TATGCAAGAAAAAGTATATTTTAAATGAAATATCATTAATGATGTTAATCACAGG	540
Db	18280	TATGCAAGAAAAAGTATATTTTAAATGAAATATCATTAATGATGTTAATCACAGG	18221
Qy	541	CTTCACTGTAATAAATAAATTTTAAAAACCAATCAATGSGTGGTTTTGGGTACACA	600
Db	18220	CTTCACTGTAATAAATAAATTTTAAAAACCAATCAATGSGTGGTTTTGGGTACACA	18161
Qy	601	AGTAAGGAGCCCACTGAAGAAACGGTCCGACTGTCTCTCTCTTTCTTTCTCTGTA	660
Db	18160	AGTAAGGAGCCCACTGAAGAAACGGTCCGACTGTCTCTCTCTTTCTTTCTCTGTA	18101
Qy	661	TTATTTGGCTACTACTATTTTACATCTCACAGAAAAAAAAGATTGAGAACTAAAC	720
Db	18100	TTATTTGGCTACTACTATTTTACATCTCACAGAAAAAAAAGATTGAGAACTAAAC	18041
Qy	721	ACAAGTTACTAAGGAGTATATCTCTGACCTTAACATCACGCTTTATATTCACCTCA	780
Db	18040	ACAAGTTACTAAGGAGTATATCTCTGACCTTAACATCACGCTTTATATTCACCTCA	17981
Qy	781	CGTAACTCACCCCTCCAAAAACATGTAATTACAGTGGACGATGATCACAGAACA	840
Db	17980	CGTAACTCACCCCTCCAAAAACATGTAATTACAGTGGACGATGATCACAGAACA	17921
Qy	841	TGATTTCTTAATATGAACTCAATGTACTTGAACACACACGACCCCAATTTTACATTGA	900
Db	17920	TGATTTCTTAATATGAACTCAATGTACTTGAACACACACGACCCCAATTTTACATTGA	17861
Qy	901	TGAAAAAATATATATTTTGTGGAGAAAGAAAGAGATTTCTTCTTCCATTCACAG	960
Db	17860	TGAAAAAATATATATTTTGTGGAGAAAGAAAGAGATTTCTTCTTCCATTCACAG	17801
Qy	961	GAAAGAAAAGGTATTCCTCGAGACATTAATCTCACCTCCCTCTCTCTCTTCAAT	1020
Db	17800	GAAAGAAAAGGTATTCCTCGAGACATTAATCTCACCTCCCTCTCTCTCTTCAAT	17741
Qy	1021	CAGTCTACGTTACACAACTTTTCAACCCATTTCAAGGCTCTCCGGAAGTTTCGAG	1080
Db	17740	CAGTCTACGTTACACAACTTTTCAACCCATTTCAAGGCTCTCCGGAAGTTTCGAG	17681
Qy	1081	GAGTGTGTTGTTGTTTTCCGATGACTTCGATGAGACTAGTCGACATCAGACAGTCG	1140

D	b	17680	GGGTTGGTTGTTGGTTTTCCGATGACTTCGGATGAGCTACGTCACATCAGCAGCTGC	17621
O	y	1141	AGTCGCGCGGGGCGAGCAGCCGCGAGAGGAAGCCGTCGTGAGAGAGAAAGGAGATTA	1200
D	b	17620	AGCTCGCGCGGGGCGAGCAGCGCGCGAGGAGAGCGTCGTGAGAGAGAAAGGAGATTA	17561
O	y	1201	TCCGAGAGAGAAACGAGAGAAAGCCTAGCTGCGAAGATATPACCTAGGCTTAAGC	1260
D	b	17560	TCCGAGAGAGAGAAACGAGAGAGCCTAGCTGCGAAGATATPACCTAGGCTTAAGC	17501
O	y	1261	TCAAGGTGATTAATTAATTTGCTTAACATTTGTGATTAATAATGAAGTCCCTTAAGCTCTT	1320
D	b	17500	TCAAGGTGATTAATTAATTTGCTTAACATTTGTGATTAATAATGAAGTCCCTTAAGCTCTT	17441
O	y	1321	TGTTGAAGCTGTGGGTTGTTGAGAGAGATGCTACTATTAATGCAAGTGAAGACTTT	1380
D	b	17440	TGTTGAAGCTGTGGGTTGTTGAGAGAGATGCTACTATTAATGCAAGTGAAGACTTT	17381
O	y	1381	CTCCATTTTTTCAGATCTGAGCGTGTATTAATGATGTTTTGATGTTGAATCTGAAT	1440
D	b	17380	CTCCATTTTTTCAGATCTGAGCGTGTATTAATGATGTTTTGATGTTGAATCTGAAT	17321
O	y	1441	CGTTGATTTCAATTTGGTTTAAATGGGTTGAATCTGAGATTTGAAGGTTTTCTCAAG	1500
D	b	17320	CGTTGATTTCAATTTGGTTTAAATGGGTTGAATCTGAGATTTGAAGGTTTTCTCAAG	17261
O	y	1501	TGAAATTTGAATCATCAGAAACATATGATGATCTGATTTCTCAAGTGAATTTATGGTT	1560
D	b	17260	TGAAATTTGAATCATCAGAAACATATGATGATCTGATTTCTCAAGTGAATTTATGGTT	17201
O	y	1561	TTCTTTCTAATTTTAAGTTATTAATGGTATATGCTAAAGTCTTAATCTTTAATGATGATA	1620
D	b	17200	TTCTTTCTAATTTTAAGTTATTAATGGTATATGCTAAAGTCTTAATCTTTAATGATGATA	17141
O	y	1621	CTTGGTCAAAGTCATGTCATGTTGTTCTTTTGCTTAACCTGATGATGATGTTTGA	1680
D	b	17140	CTTGGTCAAAGTCATGTCATGTTGTTCTTTTGCTTAACCTGATGATGATGTTTGA	17081
O	y	1681	TTGGTATTTGTTTGGCTTTGTTGAGATATCAGGATGCAAGCCTTAACCTGTGAGATA	1740
D	b	17080	TTGGTATTTGTTTGGCTTTGTTGAGATATCAGGATGCAAGCCTTAACCTGTGAGATA	17021
O	y	1741	GCTGGGACTTCATCTCGAGTATCTCCATTTATCATCAGAAACGAGGCCCTCTTCATCA	1800
D	b	17020	GCTGGGACTTCATCTCGAGTATCTCCATTTATCATCAGAAACGAGGCCCTCTTCATCA	16961
O	y	1801	GCTTTTCAAAGTCCCATCCCATCTTACCAAGTTAGCCGCTCTTCATCATCTCCGAGT	1860
D	b	16960	GCTTTTCAAAGTCCCATCCCATCTTACCAAGTTAGCCGCTCTTCATCATCTCCGAGT	16901
O	y	1861	CTTTTCGGGATGAAACCAATTAACAATGTCCTCTACATTTCTCCCTTCTCAGAAAT	1920
D	b	16900	CTTTTCGGGATGAAACCAATTAACAATGTCCTCTACATTTCTCCCTTCTCAGAAAT	16841
O	y	1921	GATGGCATTCCTTCTTCTCTTCTCTCTCCCTCAGAAATCTCAACAGTTGTCCAGTTACCCCA	1980
D	b	16840	GATGGCATTCCTTCTTCTCTTCTCTCTCTCCCTCAGAAATCTCAACAGTTGTCCAGTTACCCCA	16781
O	y	1981	CCGCTCTCATGCGGACTTCTTAACAAACCGGAAACCGTTGCCCTAACCTGGAAATCTATCGCT	2040
D	b	16780	CCGCTCTCATGCGGACTTCTTAACAAACCGGAAACCGTTGCCCTAACCTGGAAATCTATCGCT	16721
O	y	2041	AAGCAATCAACGCGCATTTGCTAAACAAATCAATGCGCTCTTTTAATTAATCTTTCTATGCG	2100
D	b	16720	AAGCAATCAACGCGCATTTGCTAAACAAATCAATGCGCTCTTTTAATTAATCTTTCTATGCG	16661
O	y	2101	GTTTCTGCACTGTAGTCCGACATATGCGCAATGCGCAATGTTTCATACCTGCGTACTATACCT	2160
D	b	16660	GTTTCTGCACTGTAGTCCGACATATGCGCAATGCGCAATGTTTCATACCTGCGTACTATACCT	16601
O	y	2161	GAATGTGATGCTGACTCTTCCACTGTGATTTCTGTGATTTGATGATTAAGCTTTCAGAA	2220


```

Db      1660  GAATGTGATGAGTCTGACTCTTCCACTGTGATTCGTGATTCATGATAGTTCAGAG 16541
Oy      2221  TTGGCAACAACACGACCATTTCTGCTCTATGTCGCAACCTCTTCACTTCAATCTT 2280
Db      16540  TTTCGACAAACAACAGCATCTTCTGCTCTATGTCGCAACCTCTTCACTTCAATCTT 16441
Oy      2281  GTGAACCTGCGCCCTCAGAGATGTCCTCAAAATCTGCTGCTTCCAGAGATGTCGCA 2340
Db      16480  GTGAACCTGCGCCCTCAGAGATGTCCTCAAAATCTGCTGCTTCCAGAGATGTCGCA 16421
Oy      2341  AGCTGTGAGTTAAATTTGAGAAATGACCAAGTTAAACCTGGAAGAGAGAGATCAT 2400
Db      16420  AGCTGTGAGTTAAATTTGAGAAATGACCAAGTTAAACCTGGAAGAGAGAGATCAT 16361
Oy      2401  GATGTGGTATGAGAGATCTTGAGCTTACACTTGGAAATGGAAGGCTGCTGAGCAT 2460
Db      16360  GATGTGGTATGAGAGATCTTGAGCTTACACTTGGAAATGGAAGGCTGCTGAGCAT 16301
Oy      2461  AAACAACCTAGGCAACCCAAATGCGATGTCATGGAATGGAAGAACTAATCTCTTGAG 2520
Db      16300  AAACAACCTAGGCAACCCAAATGCGATGTCATGGAATGGAAGAACTAATCTCTTGAG 16241
Oy      2521  TATTTCTCTTTCGCGCAGATTTGGATCTTATGATTCATGATCTTCTTCACTTAT 2580
Db      16240  TATTTCTCTTTCGCGCAGATTTGGATCTTATGATTCATGATCTTCTTCACTTAT 16181
Oy      2581  TATCCAAATATGCTGCGCAAGCCTTCTCCATGGAAGATGGAAGTGTAGAGTGTAT 2640
Db      16180  TATCCAAATATGCTGCGCAAGCCTTCTCCATGGAAGATGGAAGTGTAGAGTGTAT 16121
Oy      2641  TCAATTCATGAATTTGGTTTCAAAAGCATTAATTTGTATGATTAATAA 2687
Db      16120  TCAATTCATGAATTTGGTTTCAAAAGCATTAATTTGTATGATTAATAA 16074

RESULT 3
AY065049      1476 bp  mRNA  linear  PLAN 07-JAN-2002
LOCUS      Arabidopsis thaliana At1g75080/F9E10.7 mRNA, complete cds.
ACCESSION  AY065049
VERSION     AY065049.1 GI:18086460
KEYWORDS
SOURCE      FLI CDNA
ORGANISM    Arabidopsis thaliana (thale cress)
REFERENCE   Arabidopsis thaliana
AUTHORS     Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bower,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
            Setou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
            Arabidopsis cDNA clones
            Unpublished
            2 (bases 1 to 1476)
TITLE       Arabidopsis thaliana
JOURNAL
REFERENCE
AUTHORS

```

```

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Saitou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shim,P., Banh,J., Bower,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
    source
        1..1476
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:3702"
            /chromosome="1"
            /clone="RAFL04-20-E20(R21167)"
            /note="ecotype: Columbia"
            1..205
            206..1216
            /note="unknown protein"
            /codon_start=1
            /product="At1g75080/F9E10.7"
            /protein_id="AAL57684.1"
            /db_xref="GI:18086461"
            /translation="MTSDGATSTSAAAAAAAAAARRKPSRRENNRRRRRAVA
            AKITVGRAGDINLPKHCONNEVLKLCVAGAVVEDGTYRKCKPLPGLAGTS
            SRVTPSSQSGPLSPAFQSPISPVQVSPSPSPSPSPSPSPSPSPSPSPSPSP
            IPSLPLRLSNSCPVTPPLSPSPSKPKPLPNMESIAKOSMAIAKOSMAFENVFVA
            VSAAPSTHROHPATIPEDCSDSSTVDGSHWISFQKPAQOQPSASVPTSPPT
            NLVKNPAQOKSPNTAAFOELIGQSEPFENSVQVPMWGERIHVDGMDLELTNGKA
            RG"

BASE COUNT      390 a 1217 . 1476
ORIGIN
3' UTR
Query Match      40.5%; Score 1088.8; DB 8; Length 1476;
Best Local Similarity 80.7%; Pred. No. 3; E-220;
Matches 1444; Conservative 0; Mismatches 2; Indels 344; Gaps 1;

Oy      898  AGATGAAAAAATATTTATTTGTTGGAGAAGAAAGAGATTCCTTCTGATTC 957
Db      1  AGATGAAAAAATATTTATTTGTTGGAGAAGAAAGAGATTCCTTCTGATTC 60
Oy      958  AGGAGAGAAAGCGATTCCTGCGAGCACTTAATCTTCACTCTCTCTTCTTCTT 1017
Db      61  AGGAGAGAAAGCGATTCCTGCGAGCACTTAATCTTCACTCTCTCTTCTTCTT 120
Oy      1018  CATTAGCTACGTTCCACAGATCTTCAACCACTTCAAGCTCTCTCGGAAGTTTC 1077
Db      121  CATTAGCTACGTTCCACAGATCTTCAACCACTTCAAGCTCTCTCGGAAGTTTC 180
Oy      1078  GAGGGGTGTTGTTGTTTCCCGATGACTTGGATGAGCTGCTGAGCATGAGC 1137
Db      181  GAGGGGTGTTGTTGTTTCCCGATGACTTGGATGAGCTGCTGAGCATGAGC 240
Oy      1138  TGCAGCTGCGCGCGCGAGCAAGCGCGAGAGAGAACCGTCTGAGAGAAAGAGAA 1197
Db      241  TGCAGCTGCGCGCGCGAGCAAGCGCGAGAGAGAACCGTCTGAGAGAAAGAGAA 300
Oy      1198  TAATCGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
Db      301  TAATCGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Oy      1258  AGCTCAAGGTATTAATTTGCTTAAACATGTAATTAATGAGAGCTCTTAAGCT 1317

```

COMMENT

TITLE
JOURNAL
REFERENCE
AUTHORS

COMMENT

Db 361 AGCTCAGGTATTAATTTGCTTAAACATTTGTGATTAATGAAGTCTTAAAGTCT 420
 Oy 1318 TTGTGTTGAAGCTGTTGGTTGTTGAAGAGATGTTACTTATTCGCAAGGTGAAGAC 1377
 Db 421 TTGTGTTGAAGCTGTTGGTTGTTGAAGAGATGTTACTTATTCGCAAGGTGAAGAC 470
 Oy 1378 TTTCTTCATTTTTCAGATCTGAGCTGTTTATTGATGTTTGTGATGTTGAATCTGA 1437
 Db 471 ----- 470
 Oy 1438 ATTGTTGATTTGATTTGTTGTTAAATGGTTGAAATCTGAGATTTGAGGTTTCTCA 1497
 Db 471 ----- 470
 Oy 1498 AAGTGAATTTGAATCATCAGAAAATAATGATGATCTGATTTCTCAAGTGAATTTATGG 1557
 Db 471 ----- 470
 Oy 1558 GTTTTCTTCTAATTTTAGATTAATTTGATGTTAAGTCTTAATCTTTATGATGATG 1617
 Db 471 ----- 470
 Oy 1618 ATACTGTGTCAGAAATGATTCATGATGTTGTTCTTTGCTTACCTGATGATGATGT 1677
 Db 471 ----- 470
 Oy 1678 TGATGTTATTTGTTTGTCTTTTGTGGAGTATCAGGATGCAAGCTTTACCTGTGAG 1737
 Db 471 ----- 470
 Oy 1738 ATAGCTGGAGCTTATCTGAGTATCTGATTCATTCATCAGAAACCAAGAGCCCTCTTCA 1797
 Db 497 ATAGCTGGAGCTTATCTGAGTATCTGATTCATTCATCAGAAACCAAGAGCCCTCTTCA 556
 Oy 1798 TCAGCTTTAAAGTCCCATCCCATCTTACCAAGTTGACCGGTTCTTATCATCTCCG 1857
 Db 557 TCAGCTTTAAAGTCCCATCCCATCTTACCAAGTTGACCGGTTCTTATCATCTCCG 616
 Oy 1858 AGTCTCTGCGGTTGAACCAATAAACAATGCTCTCATCTTCCCTTCTCTGAGA 1917
 Db 617 AGTCTCTGCGGTTGAACCAATAAACAATGCTCTCATCTTCCCTTCTCTGAGA 676
 Oy 1918 AATGTGGCATCTCTTCT 1977
 Db 677 AATGTGGCATCTCTTCT 736
 Oy 1978 CCACCGGTCTCATGCGGATCTCTAAGAACCCGAAACCGTTGCTTAATCTGGATCTATC 2037
 Db 737 CCACCGGTCTCATGCGGATCTCTAAGAACCCGAAACCGTTGCTTAATCTGGATCTATC 796
 Oy 2038 GCTAAGCAATTCATGCGGATCTCTAAGAACCAATGAGGCTCTTTTAATCTCTCTCTAT 2097
 Db 797 GCTAAGCAATTCATGCGGATCTCTAAGAACCAATGAGGCTCTTTTAATCTCTCTCTAT 856
 Oy 2098 GCGGTTTCTGACCTGCTAGTCCGACATGCGGACCAAGTTTCATACCTGCTACTATA 2157
 Db 857 GCGGTTTCTGACCTGCTAGTCCGACATGCGGACCAAGTTTCATACCTGCTACTATA 916
 Oy 2158 CTTGAATGTGATGCTGATCTCTTCCACTGTTGATTTGCTCATGATGATGATGATGATG 2217
 Db 917 CTTGAATGTGATGCTGATCTCTTCCACTGTTGATTTGCTCATGATGATGATGATGATG 976
 Oy 2218 AAGTTTGACAAACAGGCAATTCCTGCTCTAATGAGTCCCAAGCTCTCTACTCTCAAT 2277
 Db 977 AAGTTTGACAAACAGGCAATTCCTGCTCTAATGAGTCCCAAGCTCTCTACTCTCAAT 1036
 Oy 2278 CTTGTTGAACCTGCGCTCAGAGATGCTCTCAAAATGCTGCTCTCTCAAGAGATGCT 2337
 Db 1037 CTTGTTGAACCTGCGCTCAGAGATGCTCTCAAAATGCTGCTCTCTCAAGAGATGCT 1096
 Oy 2338 CAAAGCTGAGTTTAAATTTGAGAAATGCGCAAGTTAAACCTTGGGAAAGAGAGAGATA 2397
 Db 1097 CAAAGCTGAGTTTAAATTTGAGAAATGCGCAAGTTAAACCTTGGGAAAGAGAGAGATA 1156

Oy 2398 CATGATGTGGATATGAGAGATCTTGAGCTTACACTTGSAAATGGAAGGCTCGTGTTGA 2457
 Db 1157 CATGATGTGGATATGAGAGATCTTGAGCTTACACTTGSAAATGGAAGGCTCGTGTTGA 1216
 Oy 2458 CATTAACAACTAGGCAAAACCAAAATGAGCATGATGATGATGATGATGATGATGATGATG 2517
 Db 1217 CATTAACAACTAGGCAAAACCAAAATGAGCATGATGATGATGATGATGATGATGATGATG 1276
 Oy 2518 GAGTATTTTCTTCTGCTCCAGGATTTGGATCTTTATGGAATCTCATGATGATGATGATG 2577
 Db 1277 GAGTATTTTCTTCTGCTCCAGGATTTGGATCTTTATGGAATCTCATGATGATGATGATG 1336
 Oy 2578 TATTAACCAAAATGAGTCCCAAGGCTTCTCCATGGAAGCATGAGTGTGAGTGTG 2637
 Db 1337 TATTAACCAAAATGAGTCCCAAGGCTTCTCCATGGAAGCATGAGTGTGAGTGTG 1396
 Oy 2638 TATTAACCAAAATGAGTCCCAAGGCTTCTCCATGGAAGCATGAGTGTGAGTGTG 2697
 Db 1397 TATTAACCAAAATGAGTCCCAAGGCTTCTCCATGGAAGCATGAGTGTGATGATGATG 1446

RESULT 4
 LOCUS AY087257 1501 bp mRNA linear PLN 14-APR-2003
 DEFINITION Arabidopsis thaliana clone 33367 mRNA, complete sequence.
 ACCESSION AY087257
 VERSION AY087257.1 GI:21405981
 KEYWORDS FLI, CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS Haas, B.J., Volkov, N., Town, C.D., Troukhan, M., Alexandrov, N.,
 Feldmann, K.A., Flavell, R.B., White, D., and Salzberg, S.L.
 TITLE Full-length messenger RNA sequences greatly improve genome
 annotation
 JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
 MEDLINE 22088475
 PUBMED 12093376

REFERENCE
 AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
 Feldmann, K.
 TITLE Full-length cDNA from Arabidopsis thaliana
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1501)
 AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
 Feldmann, K.

COMMENT
 JOURNAL Direct Submission
 Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
 Malibu, CA 90265, USA
 This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to RIKR and Genbank. The following quality assessment of
 this set was done by comparison with known proteins: two percent of
 the clones are estimated to be 5'-truncated; less than one percent
 are 3'-truncated; approximately two percent represent alternative
 splice variants, including unspliced introns and spliced exons; one
 percent may contain premature stop codons; five percent may have
 frame shifts in a coding region. A sequence is considered to be
 5'-truncated if it lacks the translation initiation start (ATG). A
 sequence is considered to be 3'-truncated if it lacks the
 C-terminal end of the encoded protein. Please note that these cDNA
 sequences are derived from the Ws or Ls ecotypes and therefore
 may contain polymorphisms when compared to sequences from Col-0.
 Genet carried out the library production and sequencing of the
 full-length clones. Ceres, Inc. carried out the clustering of the
 5' sequences, selection of clones, and sequence assembly.

FEATURES
 source

1. 1501
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"

```

/db_xref="taxon:3702"
/clone="33367"
187..1197
/codon_start=1
/product="unknown"
/protein_id="AA064812.1"
/db_xref="GI:21592862"
/translation="MTSDGATSTAAAAAARRKPSRERENRRRRRAVA
AKIYTLGAOQDYLPHKDNNEVLKALCVAGWVEEDGTYRKGKPLGELAGTS
SRVTPSSONOSPLSAPSPBIPSYOVSPSSPSPSPSPSPSPSPSPSPSPSP
IPSLPBLRISNSCPYTPVPSPTSPKPKPLPNHESIAKOSMAIAKOSMSFNTFYA
VSAPASPTLRHOFHPATIPEDCDSSTVDSGWISFQKFAQQQPSASAVPTSPF
NLVPAEQWSPNTAFQBIQSSSEFKENSQVPMWEGRIHDVEMDLIELLQNGRA
RG"
BASE COUNT      389 a      349 c      312 g      451 t
ORIGIN
Query Match      39.8% Score 1068.2; DB 8; Length 1501;
Best Local Similarity 80.4%; Pred. No. 8,7e-216;
Matches 1424; Conservative 0; Mismatches 3; Indels 344; Gaps 1;

Oy 917 ATTTGTTGAGAGAGAGAGAGATTTCTTTCTTCGATTCAGGAGAGAGAGAGAGATTT 976
Db 1 ATTTGTTGAGAGAGAGAGAGATTTCTTTCTTCGATTCAGGAGAGAGAGAGAGATTT 60

Oy 977 CCTGTGAGAGACTACTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1036
Db 61 CCTGTGAGAGACTACTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120

Oy 1037 AATCTTTACCCCACTATTCAAAAGCTCTCTCCGAGATTTCAGAGGGTGGTGGTGGT 1096
Db 121 AATCTTTACCCCACTATTCAAAAGCTCTCTCCGAGATTTCAGAGGGTGGTGGTGGT 180

Oy 1097 TTCCCATGACTTCGATGAGAGTCACTGACATCAGACGTCGACGTCGCGCGCGCA 1156
Db 181 TTCCCATGACTTCGATGAGAGTCACTGACATCAGACGTCGACGTCGCGCGCGCA 240

Oy 1157 GCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1216
Db 241 GCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

Oy 1217 CGGAGAGAGAGCTGTAGCTGCGAATATACATCGGCTTAGAGCTCAAGGTGATTAAT 1276
Db 301 CGGAGAGAGAGCTGTAGCTGCGAATATACATCGGCTTAGAGCTCAAGGTGATTAAT 360

Oy 1277 TTGCTTAAACATTTGTATATAATGAAGTCTTTAAAGCTTTTGTGTGAAGCTGGTGG 1336
Db 361 TTGCTTAAACATTTGTATATAATGAAGTCTTTAAAGCTTTTGTGTGAAGCTGGTGG 420

Oy 1337 GTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1396
Db 421 GTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451

Oy 1397 TGTGAGCTGTTTATGATGTTTGTGATGTTGAATCTGAATCTGATTTCAATTTGT 1456
Db 452 TGTGAGCTGTTTATGATGTTTGTGATGTTGAATCTGATTTCAATTTGT 451

Oy 1457 GGTAAATGGGTTTGAATCTGAGAAATTTGAGGGTTTTCGAAAGTGAATTTGAATCA 1516
Db 452 GGTAAATGGGTTTGAATCTGAGAAATTTGAGGGTTTTCGAAAGTGAATTTGAATCA 451

Oy 1517 GAAACTATGATGATCTGATTTCTCAAGTGAATTTATGGTCTTTCTTAATTTTAG 1576
Db 452 GAAACTATGATGATCTGATTTCTCAAGTGAATTTATGGTCTTTCTTAATTTTAG 451

Oy 1577 AGTTATTTGATGCTAAAGCTTAATCTTTTATGATGATGATGATGATGATGATGAT 1636
Db 452 AGTTATTTGATGCTAAAGCTTAATCTTTTATGATGATGATGATGATGATGATGAT 451

Oy 1637 TGCATTTGTTCTTTCTTCTTACCTGTGATGATGATGATGATGATGATGATGATG 1696
Db 452 TGCATTTGTTCTTTCTTCTTACCTGTGATGATGATGATGATGATGATGATGATG 451

```

```

Oy 1697 TTTTGTGAGATCAGGAGATGCAAGCTTTTACCTGGAGATGATGATGATGATGATGAT 1756
Db 452 TTTTGTGAGATCAGGAGATGCAAGCTTTTACCTGGAGATGATGATGATGATGATGAT 496

Oy 1757 GAGTAACTCATATTTATTCACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1816
Db 497 GAGTAACTCATATTTATTCACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556

Oy 1817 TCCCATCTTACCAAGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1876
Db 557 TCCCATCTTACCAAGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616

Oy 1877 CAATTAACAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1936
Db 617 CAATTAACAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 676

Oy 1937 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1996
Db 677 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 736

Oy 1997 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2056
Db 737 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 796

Oy 2057 TTGCTTAAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2116
Db 797 TTGCTTAAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 856

Oy 2117 GTCCGACACATGCGCACAGAGTTTATACACCTGCTACTAATCCGAATGATGATGATG 2176
Db 857 GTCCGACACATGCGCACAGAGTTTATACACCTGCTACTAATCCGAATGATGATGATG 916

Oy 2177 ACTCTTCACTCTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2236
Db 917 ACTCTTCACTCTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 976

Oy 2237 CATTTCTGCTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2296
Db 977 CATTTCTGCTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1036

Oy 2297 AGCAGATGCTGCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2356
Db 1037 AGCAGATGCTGCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1096

Oy 2357 TTGAGAAATGCGCAAGTTAAACCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2416
Db 1097 TTGAGAAATGCGCAAGTTAAACCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1156

Oy 2417 ATCTTGAAGTTACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2476
Db 1157 ATCTTGAAGTTACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1216

Oy 2477 CCAAAATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2536
Db 1217 CCAAAATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276

Oy 2537 CAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2596
Db 1277 CAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1336

Oy 2597 CCAAAAGCTTTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2656
Db 1337 CCAAAAGCTTTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1396

Oy 2657 GTTCAAAAAGCATTTTGTAGATAAAAAA 2687
Db 1397 GTTCAAAAAGCATTTTGTAGATAAAAAA 1427

```

RESULT 5
 AY093747
 LOCUS
 DEFINITION Atadiopsis thailana Atig75080/75080.7 mRNA, complete cds.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
COMMENT
TITLE
JOURNAL
AUTHORS
FEATURES
SOURCE
CDS

AY093747
AY093747.1
GI:20147314
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 1011)
Shim, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Saitou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 1011)
Shim, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Saitou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (29-MAR-2002) Salk Institute Genomic Analysis Laboratory (Signal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Saitou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.
The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Shim, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bower, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.
Shim, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.
Location/Qualifiers
1..1011
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U21167"
/note="this clone is in PUNI 51 ecotype: Columbia"
1..1011
/note="unknown protein"
/codon_start=1
/product="At1g5080/F9B10.7"
/protein_id="AA10371.1"
/translation="MTSDGATSTGAAAAAARRKPSWRENNRRRRRAVA
AKITGAGAGDNLPGHNNNEVLAALCEAGWVVEDDTYRKGGKPLGELAGS
SRVTYSSNONGSPSLSAFQSPRIEYQVSSPSRSGEPNNNNSTFPPLRNG
IPSLPLRLSNSCPYPPISSPKPKPLPNMESIAKSMIAKSMASPNPFYA
VSAASPTHHQHTPATPTPECDSSTSDSWSWIFQKRAQOQPSASAMVPSPTF
NLVNPAPQOMSPTNPAFOEIGOSSEFPENSQVAFWEGERLHDVGMEDLETLANGNA
RG"

BASE COUNT 261 a 261 c 226 g 273 t
ORIGIN
Query Match 27.6%; Score 742.8; DB 8; Length 1011;
Best Local Similarity 99.7%; Pred. No. 8.3e-147;
Matches 744; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1712 AGGATGCAAGCCCTTATCCTGAGATAGTGGAGACTTCATCTGATTAATTCATATT 1771
|||
266 AGGATGCAAGCCCTTATCCTGAGATAGTGGAGACTTCATCTGATTAATTCATATT 325
|||
1772 CATCACAGAACAGAGCCCTTTCATCAGCTTTCAAGATCCATTCATTCACCAAG 1831
|||
326 CATCACAGAACAGAGCCCTTTCATCAGCTTTCAAGATCCATTCATTCACCAAG 385
|||
1832 TTAGCCGCTCTTTCATCATATTCGCGAGTCTTTCGCGGAGAACCAATTAACAATGT 1891
|||
386 TTAGCCGCTCTTTCATCATATTCGCGAGTCTTTCGCGGAGAACCAATTAACAATGT 445
|||
1892 CCTGACATTCCTCCCTTCCAGAAATGAGTGCATTCCTCTCTCCCTCCCTCA 1951
|||
446 CCTGACATTCCTCCCTTCCAGAAATGAGTGCATTCCTCTCTCCCTCCCTCA 505
|||
1952 GAATCTCAAGAGTTGTCAGTTACCCGAGCTTCATGCGGACCTTCAAGACCGCA 2011
|||
506 GAATCTCAAGAGTTGTCAGTTACCCGAGCTTCATGCGGACCTTCAAGACCGCA 565
|||
2012 AACGCTGCTTACGAGATATTCGCTTACCAATCCATGCGCATTTGCTTAACATCA 2071
|||
566 AACGCTGCTTACGAGATATTCGCTTACCAATCCATGCGCATTTGCTTAACATCA 625
|||
2072 TGGCGCTTTTAAATATCTTCTTAAAGGTTTTCAGACCTGCTAGTCCGACATGCGC 2131
|||
626 TGGCGCTTTTAAATATCTTCTTAAAGGTTTTCAGACCTGCTAGTCCGACATGCGC 685
|||
2132 ACCAGTTTCATACCTGCTTACATATACCTGATGATGATGATGATGATGATGATG 2191
|||
686 ACCAGTTTCATACCTGCTTACATATACCTGATGATGATGATGATGATGATGATG 745
|||
2192 ATTCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2251
|||
746 ATTCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 805
|||
2252 TGGTGGCAACCT 2311
|||
806 TGGTGGCAACCT 865
|||
2312 ATACTGCTGCT 2371
|||
866 ATACTGCTGCT 925
|||
2372 TTAAACCTGGAGAGAGAGATATCATGATGATGATGATGATGATGATGATGATG 2431
|||
926 TTAAACCTGGAGAGAGAGATATCATGATGATGATGATGATGATGATGATGATG 985
|||
2432 TTGGAATGGAGAGGCTGCTGCTGA 2457
|||
986 TTGGAATGGAGAGGCTGCTGCTGA 1011
|||

RESULT 6
AC025808 120977 bp DNA linear PLN 11-OCT-2000
LOCUS AC025808
DEFINITION Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome I, complete sequence.
ACCESSION AC025808
VERSION AC025808.8 GI:7636235
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE AUTHORS	1 (bases 1 to 120977) Shin,S., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,Q., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Festerspiel,N.A., Theologis,A. and Ecker,J.R. Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome I	CDS	NVAEARKLEKMLEKGLVSGDKQGVDIIPKEVVLDPENVKLGSTGVLENNSELDYDR. RUSEAVVY" complement(4647..5702) /note="putative glycosyltransferase-like protein emb CAB42905.1; similar to EST gb A1998490.1" /codon_start=1 /evidence=not experimental /product="F18014.2" /protein_id="AAF79456.1" /db_xref="GI:8778448" /translation="MSQHLILLLSLLHLKPIASATTIOKEKAPOFYNADCPIL DSESDDDVAKPIFCSSRAVHMTLDAVIRGSVAIVLISGSCSPNIHVPFA SASDSSSRATISSFPYLDPTVYVFNSSVSRLLSSIRSLDCCPTNARSTLAD LPCVRRVYLLSDLLVDIAKLAITLGRDVLAPECYNAFPSTFTWNSPT LSUTEDRACYPFNTGVWVILDSRMEGAVTSIEEMAMQKMRITVELSLPFLV FGLIKPVNHNQHGIGDGNFRGLCRDHPGVSLIHWGKPKMARLDAGRPCLD ALMAYDLDTQTFALDS" complement(9620..10300) /note="similar to Zn finger protein gi 6319911" /codon_start=1 /evidence=not experimental /product="F18014.3" /protein_id="AAF79455.1" /db_xref="GI:8778447" /translation="MSDVPSCSSGNDTNNDNSFEFCNTCLDAQPIVTLGHLFEW PCLYKHLHSGSKDPCVCAVIEEDRIPLTGRGSSADPRKSIPGLEVPNRPSCQ RERTQPPDPNPGFHHNGRFGFMGAPMARSARGNTLLSAFGGLIPSLNHLHF GPPDAMYGAAASGGRPHGFSNPFHGSJHMSYQRHTROGQODHLRLILVFEV VFSFPLS" join(12523..12577,12686..13374) /note="hypothetical protein" /codon_start=1 /evidence=not experimental /product="F18014.4" /protein_id="AAF79420.1" /db_xref="GI:8778412" /translation="MAISTSHLPISTIARTISVSSTFTTLTHGCTTPPGILT ANGADLDGGRPLAGSSVTFTVSPMSGRFARTYCNFDAGSGKCGTDCSSKIK AGAGAPPEATLAEFTIGSGKNAVQDFVSLVDQVYQMTITPGGSGCTTACVY SDVNAICPEKELQVTPSGVACKSACEAFNKECYCCTGAVSTPATCPPTNYSKIFQXA CPSAYSYVDDSSFTCTNAYEISFCS" complement(join(14020..14136,14236..14304,14402..14464, 14557..14666,14879..15005,16274..16486)) /note="unknown protein; similar to EST gb H76493.1" /codon_start=1 /evidence=not experimental /product="F18014.5" /protein_id="AAF79454.1" /db_xref="GI:8778446" /translation="WLEAVDSSGVNGGEPQIOFYGDCSSSEELSLPRHTKVVTG NNRTSVVLGLOGVVRKAVGLGWMVLVTNGIEVLAQRNALSVLEPPGNEDDDL FENTGRNGSDMTSEDTLKPHKSKLRGSSSRSHMTSRSLSDSSGKSGFPPPEM KVDLSKLEMPALINVRHNLVDALPNSKEGLDIDVQHFPMSQMDDELQVYVGFQA AKRMKACKPOSKESSRNTDNCIS" join(17994..18489,18568..18696,18814..18847,19006..19099, 19170..19347,19441..19479,19616..19715,19855..19994) /note="hypothetical protein" /codon_start=1 /evidence=not experimental /product="F18014.6" /protein_id="AAF79421.1" /db_xref="GI:8778413" /translation="MAKTDKLAQFLDSQIYESDEFNPFLLDTVRIITNRSYTRFKVSP AYSRFPNSKOLNCHSSSESNPKRKRKOKNSFFHLPSVGOASNLHOCARLFLSAH ESFLKEILLSTKLSDNDDDSLNKKCCDDSEVFEIGGWQAPEVETLISFL HCDNGESCNORVQVFNNTVYVNEIGSEVSEFNRYIMRNCFCVYSDLHITNTL VPAISEBYNLIIVDPENNSAHSKSFVITYLMTPLPNOFLFLPIQLHABEAL VALMTNKEKLSFEVERLEFPWGIKIVATVMTVQKPDGTLICDLDLAGEKRSDF KLAKQOIMSLIPGDFSRPKPIGDLILHTPDSOPARCHELFPAREMAAGMTSGNEPL HFDSRFLYK" join(42037..22300,22618..23361) /note="unknown protein; similar to ESTs gb A1992723.1, gb t41777.1, dbj av440980.1, gb A1997341.1, and
REFERENCE AUTHORS	2 (bases 1 to 120977) Ecker,J.R. Direct Submission Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA 3 (bases 1 to 120977) Ecker,J.R. Direct Submission Submitted (22-APR-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA 4 (bases 1 to 120977) Chen,R., Shin,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Davis,R., Festerspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA 5 (bases 1 to 120977) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shin,P., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Davis,R., Festerspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA On Apr 22, 2000 this sequence version replaced gi:7543634. Location/Qualifiers 1..120977 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="1" /clone="F18014" 90..2159 /note="hypothetical protein" /codon_start=1 /evidence=not experimental /product="F18014.1" /protein_id="AAF79419.1" /db_xref="GI:8778411" /translation="MISPEVSDVFCSTVYVACRSGVNDKAMPFAKETBSGLIEL NVYTNLSLNGIAMGDEGMRVRLMSERVSRRVYVTLIGYCKGLMESEAEH VVELKEKLVADQHMVGMDGCRHTQIRDAVRHDMIEIGVTKTICNSLNG YCKSQLVEAQIFSRNDWSLKPHTYNTLVDCRAGVDEALKCDQCKQEVV PLMYTNILKYSIRIAGAVDLSLKMMLKGVNADEISCTLEALFKDGDFEAV KLMBVNLAGLLDTITLNVMSGLCKMKVNAKEILDNVNIFFCKAVQVYQALSH GGVYGNLKEAFVNEWKRGIPTPIYNTLIGARFYRLNVAADVLRLARGL PTVATYTGALLITGKNCIGIDAVYITCEMIKGIITLVNICSKIANSIFLFDKIDEA CULLOKIVDFDLTPGYSLKEFLASATCTCKTIAESVNSTPKKLVPNNITVIA VAIGLCRAGLEKEDARKLFSDLSDDRFIDPEYTYTILHGCAGIDINKAFTLDEM ALKGIIPIVYTNALIKGICKGNVDRAPRLHLKLPOKGITFNALITYNTLLDGLVXSG	CDS	
REFERENCE AUTHORS	1 (bases 1 to 120977) Chen,R., Shin,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Davis,R., Festerspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA 5 (bases 1 to 120977) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shin,P., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Davis,R., Festerspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA On Apr 22, 2000 this sequence version replaced gi:7543634. Location/Qualifiers 1..120977 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="1" /clone="F18014" 90..2159 /note="hypothetical protein" /codon_start=1 /evidence=not experimental /product="F18014.1" /protein_id="AAF79419.1" /db_xref="GI:8778411" /translation="MISPEVSDVFCSTVYVACRSGVNDKAMPFAKETBSGLIEL NVYTNLSLNGIAMGDEGMRVRLMSERVSRRVYVTLIGYCKGLMESEAEH VVELKEKLVADQHMVGMDGCRHTQIRDAVRHDMIEIGVTKTICNSLNG YCKSQLVEAQIFSRNDWSLKPHTYNTLVDCRAGVDEALKCDQCKQEVV PLMYTNILKYSIRIAGAVDLSLKMMLKGVNADEISCTLEALFKDGDFEAV KLMBVNLAGLLDTITLNVMSGLCKMKVNAKEILDNVNIFFCKAVQVYQALSH GGVYGNLKEAFVNEWKRGIPTPIYNTLIGARFYRLNVAADVLRLARGL PTVATYTGALLITGKNCIGIDAVYITCEMIKGIITLVNICSKIANSIFLFDKIDEA CULLOKIVDFDLTPGYSLKEFLASATCTCKTIAESVNSTPKKLVPNNITVIA VAIGLCRAGLEKEDARKLFSDLSDDRFIDPEYTYTILHGCAGIDINKAFTLDEM ALKGIIPIVYTNALIKGICKGNVDRAPRLHLKLPOKGITFNALITYNTLLDGLVXSG	CDS	
REFERENCE AUTHORS	1 (bases 1 to 120977) Chen,R., Shin,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Davis,R., Festerspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA 5 (bases 1 to 120977) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shin,P., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Davis,R., Festerspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA On Apr 22, 2000 this sequence version replaced gi:7543634. Location/Qualifiers 1..120977 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="1" /clone="F18014" 90..2159 /note="hypothetical protein" /codon_start=1 /evidence=not experimental /product="F18014.1" /protein_id="AAF79419.1" /db_xref="GI:8778411" /translation="MISPEVSDVFCSTVYVACRSGVNDKAMPFAKETBSGLIEL NVYTNLSLNGIAMGDEGMRVRLMSERVSRRVYVTLIGYCKGLMESEAEH VVELKEKLVADQHMVGMDGCRHTQIRDAVRHDMIEIGVTKTICNSLNG YCKSQLVEAQIFSRNDWSLKPHTYNTLVDCRAGVDEALKCDQCKQEVV PLMYTNILKYSIRIAGAVDLSLKMMLKGVNADEISCTLEALFKDGDFEAV KLMBVNLAGLLDTITLNVMSGLCKMKVNAKEILDNVNIFFCKAVQVYQALSH GGVYGNLKEAFVNEWKRGIPTPIYNTLIGARFYRLNVAADVLRLARGL PTVATYTGALLITGKNCIGIDAVYITCEMIKGIITLVNICSKIANSIFLFDKIDEA CULLOKIVDFDLTPGYSLKEFLASATCTCKTIAESVNSTPKKLVPNNITVIA VAIGLCRAGLEKEDARKLFSDLSDDRFIDPEYTYTILHGCAGIDINKAFTLDEM ALKGIIPIVYTNALIKGICKGNVDRAPRLHLKLPOKGITFNALITYNTLLDGLVXSG	CDS	
REFERENCE AUTHORS	1 (bases 1 to 120977) Chen,R., Shin,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Davis,R., Festerspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA 5 (bases 1 to 120977) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shin,P., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharzky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vayebey,M., Yu,G., Davis,R., Festerspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center. Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA On Apr 22, 2000 this sequence version replaced gi:7543634. Location/Qualifiers 1..120977 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="1" /clone="F18014" 90..2159 /note="hypothetical protein" /codon_start=1 /evidence=not experimental /product="F18014.1" /protein_id="AAF79419.1" /db_xref="GI:8778411" /translation="MISPEVSDVFCSTVYVACRSGVNDKAMPFAKETBSGLIEL NVYTNLSLNGIAMGDEGMRVRLMSERVSRRVYVTLIGYCKGLMESEAEH VVELKEKLVADQHMVGMDGCRHTQIRDAVRHDMIEIGVTKTICNSLNG YCKSQLVEAQIFSRNDWSLKPHTYNTLVDCRAGVDEALKCDQCKQEVV PLMYTNILKYSIRIAGAVDLSLKMMLKGVNADEISCTLEALFKDGDFEAV KLMBVNLAGLLDTITLNVMSGLCKMKVNAKEILDNVNIFFCKAVQVYQALSH GGVYGNLKEAFVNEWKRGIPTPIYNTLIGARFYRLNVAADVLRLARGL PTVATYTGALLITGKNCIGIDAVYITCEMIKGIITLVNICSKIANSIFLFDKIDEA CULLOKIVDFDLTPGYSLKEFLASATCTCKTIAESVNSTPKKLVPNNITVIA VAIGLCRAGLEKEDARKLFSDLSDDRFIDPEYTYTILHGCAGIDINKAFTLDEM ALKGIIPIVYTNALIKGICKGNVDRAPRLHLKLPOKGITFNALITYNTLLDGLVXSG	CDS	

[illegible][illegible]

Db 1177 TC 1178

RESULT 8
LOCUS AY086340 1359 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 24157 mRNA, complete sequence.
ACCESSION AY086340
VERSION AY086340.1 GI:21405050
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1359)
Haas, B.J., Volkov, N., Town, C.D., Troupkhan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
Full-length messenger RNA sequences greatly improve genome annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
JOURNAL MEDLINE 22088475
PUBMED 12093376
2 (bases 1 to 1359)
Broyer, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1359)
Broyer, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road, Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Geneset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.
Location/Qualifiers
1..1359
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="24157"
112..1119
/codon_start=1
/product="unknown"
/protein_id="AA064408.1"
/translation="MTSDGATSTSAATAAAATRRKPSWRERENRRRRRAVAA
KITGKRAQGNVLPKCDNNEVYLKALCSAGWVEDGTYRKHPLFGDAAGSS
RATPVSHNOSPSTPDSPLSYOVSPSSSPSPERVDPHNISTIFPLANGGIP
SGSPRIASASAVTPPVSPSTRNSRPLPTWMSPTQSMASAKOSMTSINPFYV
SAPASPTHHQFAPATIPFECDSDSSTVDSGIRHQAQCOOPASASVPTPTN
LVPAPOQLSPNTAAIQEIGSSSEFKFENSQVXWBSGRIDVAMELELLTNGKRAH

BASE COUNT 356 a 317 c 285 g 401 t

ORIGIN

Query Match 19.8%; Score 532.4; DB 8; Length 1359;
Best Local Similarity 81.8%; Pred. No. 3.3e-102;

Matches 640; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

QY 1712 AGGATGCAAGGCTTACTGATGATAGCTGGACCTTCATCTGAGTAAGTCCATATT 1771
DB 374 AGGACACAGGCTCTTACTGATGATAGCTGGACCTTCATCTGAGTAAGTCCATATT 433
QY 1772 CATCAGACAGAGAGCCCTCTTCATGAGCTTCAAGTCCATCCATTCATCAAG 1831
DB 434 CTTCCTATACCAAGAGCTCTTCTTCATCTTGATAGCCCATCTTATCAAG 493
QY 1832 TTAGCCGCTCTTCATCATCTCCGAGTCTTCGCGGTGAACCAATAACACATGT 1891
DB 494 TCAGTCTCTCTCTTCATCTCCGAGTCTTCGAGTGTGATCCACAATAT-- 551
QY 1892 CCTTACATCTCTCTCTTCATCTCCGAGTCTTCGAGTGTGATCCACAATAT-- 1951
DB 552 -CTCACATCTCTCTCTTCATCTCCGAGTGTGATCCACAATAT-- 610
QY 1952 GAATCTCAAGAGTGTGATCTCCGAGTGTGATCTCCGAGTGTGATCTCCGAGTGT 2011
DB 611 GAATCTCAAGAGTGTGATCTCCGAGTGTGATCTCCGAGTGTGATCTCCGAGTGT 670
QY 2012 AACCGTGTCTTACTGGAATCTATCGCTAAGCAATCCATGGCCAT--TGCTAAACAT 2068
DB 671 AACCATGTGCTTCTGGAATCTTACCAACAAATCCATGGCTGCTTAAACAGT 730
QY 2068 CAATGCGCTCTTCTTATATCTCTTCTATGCGGTCTTCACCTGCTAGCGACATC 2128
DB 731 CAATGCTCTTCTTATATCTCTTCTATGCGGTCTTCACCTGCTAGCGACATC 790
QY 2128 GCCACAGTGTATATCTGCTGCTTACTATCTGATGATGATGATGATGATGATGAT 2188
DB 791 ATGCGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850
QY 2188 TTGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2248
DB 851 TTGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910
QY 2248 CTATGCTGCAACCTCTCTTACCTTCAATCTTGTGAACCTGCGCTCAGCATGCTC 2308
DB 911 CTATGCTGCAACCTCTCTTACCTTCAATCTTGTGAACCTGCGCTCAGCATGCTC 970
QY 2308 CAATATGCTGCTCTCTTCAAGAGTGTGATGATGATGATGATGATGATGATGATGAT 2368
DB 971 CAATATGCTGCTCTCTTCAAGAGTGTGATGATGATGATGATGATGATGATGATGAT 1030
QY 2368 AAGTTAAACCTGGAAGAGAGAGATCATGATGATGATGATGATGATGATGATGATGAT 2428
DB 1031 AAGTTAAACCTGGAAGAGAGAGATCATGATGATGATGATGATGATGATGATGATGAT 1090
QY 2428 CACTTGAATGGAAGAGCTCGTGTGATGATGATGATGATGATGATGATGATGATGAT 2488
DB 1091 CGCTTGAATGGAATGGAAGCTCGTGTGATGATGATGATGATGATGATGATGATGATGAT 1150
QY 2488 TC 2490
DB 1151 TC 1152

RESULT 9
AF134217 1257 bp mRNA linear PLN 01-JAN-2000
LOCUS AF134217 1257 bp mRNA linear PLN 01-JAN-2000
DEFINITION Arabidopsis thaliana 107 protein (107) mRNA, complete cds.
ACCESSION AF134217
VERSION AF134217.1 GI:6651068
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1257)
Kwon, H.-B.

TITLE Molecular Cloning and Characterization of a Gibberellin-responsive Gene from *Arabidopsis thaliana*

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1257)

AUTHORS Kwon, H.-B.

TITLE Direct Submission

JOURNAL Submitted (10-MAR-1999) Bioresources, National Institute of Agricultural Science and Technology, Seodun-Dong 250, Suwon, Kyunggi-do 441-707, Korea

FEATURES

source Location/Qualifiers

1.1257

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/cultiivar="Columbia"

/db_xref="taxon:3702"

1.1257

/gene="107"

/note="responsive to gibberellin in Arabidopsis mutant ga3"

55.1062

/gene="107"

/codon_start=1

/product="107 protein"

/protein_id="AF22161.1"

/db_xref="GI:6651069"

/translation="MTSDGATSTSTAAATAAATRRKREKRENRERRRRRRAVAA
KITGLAQGNINLPHKIDNNEVLKALCSAGWVEEDGTTTKRGRPLRDMAGSS
RATPYSHNQSPSTSDSPILSYQVSPSSPPSPRPVDPNHSITPFLNGGIP
SSLPRTISNAPSPVPSPTSPNPKPLTWSFTQSMMAKQSTPLNPLNPLVAV
SAPSPHROPHAPATIPEDSDSDSTVSGHSTQKFAOQOPFASVAVPSPTN
LVKAPQQLFPNTAATIGIGSSSEPKENQVFWEEGRTHDVAMDELTLNGRAH
S"

BASE COUNT 325 a 290 c 273 g 369 t

ORIGIN

Query Match 19.8%; Score 530.8; DB 8; Length 1257;
Best Local Similarity 81.7%; Pred. No. 7,2e-102;
Matches 639; Conservative 0; Mismatches 137; Indels 6; Gaps 2;

QY 1712 AGGATGCAAGCCTTACCTGGTGAATAGTGGAGATTCTCGAGTAATCCATATT 1771
DB 317 AGGAGACAAGCCTTACCTGGTGAATAGTGGAGATTCTCTCGAGCAATCCTTACT 376
QY 1772 CATCACGAAACGAGCCCTTTTCATCAGCCTTCAAAAGCCATCCATTTACCAAG 1831
DB 377 CTTCACATAACCAAGTCTCTTCTTCCACTTTGATACCCCATCTTATCTTACCAAG 436
QY 1832 TTAGCCGCTTCTTCATCATTCGCCGAGTCTTCTCGCGGGAACCAATTAACAATGT 1891
DB 437 TCAGTCTCTTCTTCTTCATTCGCCGAGTCTTCTCGAGTTGGTGAATCCACACATAT-- 494
QY 1892 CCTCTACATTTCTTCCCTTCTCGAAGATGGTGCATCTCTTCTTCTTCTTCCCTCA 1951
DB 495 -CTCCACAATCTTCCCTTCTCGAAGATGGTGAATCTCTTCACTGCTTCTTCACTTA 553
QY 1952 GAATTCACAACAGTTCAGTTACCCACCGGCTCATAGCCGCACTTCAAGACCCGA 2011
DB 554 GAATTCACAACAGTTCAGTTACCCACCGGCTCATAGCCGCACTTCAAGACCCGA 613
QY 2012 AACCGTGGCTTAATGGAATATATGCTTAAGCAATCATGAGCCAT---TGCTTAACAT 2068
DB 614 AACCATTTGCTTACTGGGAATCTTTTACCAACAATCATGCTCATGCTGCTGAACAGT 673
QY 2069 CAATGGGCTTTTAAATATATCTTCTTATGCGGTTTCTGCACTGTAAGTCCGACATC 2128
DB 674 CAATGACTTTTGAATACCCGTTTATAGGGTGTCTGCACTGCAAGTCTTCACTATC 733
QY 2129 GCCACCACTTTTCACTGCTGCTGCTATATACCTGAATGTATGAGTGAAGTCTTCCACTG 2188
DB 734 ATGCCGAGTTTCATGCTCGGCTGCTATACCTGAATGTATGAGTGAAGTGAAGTCTTCCACTG 793
QY 2189 TTGATTTGGTCAATGGATAAGCTTTAGAAAGTTTGACAACAACACCACTTCTGCGCT 2248
DB 794 TTGATTTGGTCAATGGATAAGCTTTAGAAAGTTTGACAACAACACCACTTCTGCGCT 853

QY 2249 CTATGTGCGCCACTCTCTTCACTTCAATCTTGTGAACCTGCCCTCAGACATGTCTC 2308
DB 854 CTATGTGCGCCACTCTCTTCACTTCAATCTTGTGAACCTGCCCTCAGACATGTCTC 913
QY 2309 CAATATCTGCTGCTTCCAGAGATTGTGCAAGCTCTGAGTTAAATTTGAGAAATGCC 2368
DB 914 CAATATCTGCTGCTTCCAGAGATTGTGCAAGCTCTGAGTTAAATTTGAGAAATGCC 973
QY 2369 AAGTTAAACCTCGGAGAGAGAGATCATGATGTGGTATGAGATCTTGAAGCTTA 2428
DB 974 AAGTTAAACCTCGGAGAGAGAGATCATGATGTGGTATGAGATCTTGAAGCTTA 1033
QY 2429 CACTTGAATGGAAGAGCTCTGCTTGAATTAACACTTACGCAACCAATGAGCATG 2488
DB 1034 CGCTTGAATGGAAGAGCTCTGCTTGAATTAACACTTACGCAACCAATGAGCATG 1093
QY 2489 TC 2490
DB 1094 TC 1095

RESULT 10
AF372937

LOCUS AF372937

DEFINITION Arabidopsis thaliana At1g19350/F18014_4 mRNA, complete cds.

ACCESSION AF372937

VERSION AF372937.1 GI:13937166

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 945)

Shim, P., Chen, H., Cheuk, R., Kim, C. J., Meyers, M. C., Banh, J., Bower, L., Carninci, P., Chung, M. K., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J., Pham, P. K., Quach, H. L., Sakano, H., Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G., Shinzaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.

Arabidopsis cDNA clones

Unpublished 2 (bases 1 to 945)

Shim, P., Chen, H., Cheuk, R., Kim, C. J., Meyers, M. C., Banh, J., Bower, L., Carninci, P., Chung, M. K., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J., Pham, P. K., Quach, H. L., Sakano, H., Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G., Shinzaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.

Direct Submission

Submitted (20-APR-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Shin, P., Chen, H., Cheuk, R., Kim, C. J., Koesema, B., Meyers, M. C., Traey, S. E., Banh, J., Bower, L., Chung, M. K., Goldsmith, A. D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Narusaka, M., Palm, C. J., Pham, P. K., Quach, H. L., Sakano, H., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yu, G., Davis, R. W., Theologis, A., and Ecker, J. R.

Shim, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to

COMMENT

this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/salk)
contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

1..945
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF109-13-019(R12325)"
/note="ecotype: Columbia"
1..288
289..699
/note="unknown protein"
/cdon_start=1
/product="Atg19350/F18014_4"
/protein_id="AKS0077.1"
/db_xref="GI:13937167"
/translation="MSMAKQSMSTSLNYPFVAVSAPASPTHRQFAPATIPEDCEDSD
STVDSGHMISFQKPAQOQFASMTPTSPVLVVPAPQOLSPTAAIOEIQSSEF
KPENSGVKPMESGRHIDVAMEDELTLGNKSHS"

5'UTR
CDS

3'UTR
BASE COUNT 245 a 232 c 172 g 296 t
ORIGIN

Query Match 18.7%; Score 501.4; DB 8; Length 945;
Best Local Similarity 82.0%; Pred. No. 1.2e-95;
Matches 603; Conservative 0; Mismatches 126; Indels 6; Gaps 2;

1759 GTAACATCATATTCATACAGAACAGAGCCCTCTTTCATCAGCCTTTCAGAGTCCCATC 1818
1 GCAACTCCTTACTCTCCCATACCAAGCTCTTCTTCCCATTTGATAGCCCATC 60
1819 CCATCTTACCAAGTAGCCCGCTCTTTCATCATTCCTCCAGTCTTCTCGCGGTGAACCA 1878
61 TTAATCTTACCAAGTAGCTCTTCTCTTCTTCAATCCCGAGTCTCTTCTCGAGTTGGTGAAT 120
1879 AATAACAACATGCTCTTACATCTTCCCTCTCGAAGATGGGCAATTCCTTCTCT 1938
121 CCACACAAATTT---CTCCAAATCTTCTCTCTCTCGAAGATGGGCAATTCCTTCTCT 177
1939 CTTCCTTCTCTCGAATCTTCAACAGTTGTCAGTTACCCACCGGTCTGATCGCCGACT 1998
178 CTTCCTTCTCGAATCTTCAACAGTTGTCAGTTACCCACCGGTCTGATCGCCGACT 237
1999 TCTAAGAACCCGAAACCGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2057
238 TCTAAGAACCCGAAACCGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 297
2058 --TGTCTTCAACATTCATGCGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2115
298 GCTGCTTCAACAGTTCAATGACTTCTTGAATACCCGTTTATGCGGTCTTGAACCTGCGC 357
2116 AGTCCGACATGCGCACACAGTCTTCAATACCTGCTACTTATCTTGAACCTGCGGCT 2175
358 AGTCCGACATGCGCACACAGTCTTCAATACCTGCTACTTATCTTGAATGATGATGCT 417
2176 GACTCTTCACTGTGATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2235
418 GACTCTTCACTGTGATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 477
2236 CCAATTTCTGCGCTTATGTCGCAACCTCTCTCACTTCACTTCTTGAACCTGCGGCT 2295
478 CCAATTTCTGCGCTTATGTCGCAACCTCTCTCACTTCACTTCTTGAACCTGCGGCT 537
2296 CAGCAGATGCTCTCAATACTGCTGCTTCTCAAGAGATGCTCAAGCTCTGAGTTTAA 2355
538 CAGCAGATGCTCTCAATACTGCTGCTTCTCAAGAGATGCTCAAGCTCTGAGTTTAA 597
2356 TTGTGAGAAATGAGCAAGTAAACCTTGGGAGAGAGAGATGATGATGATGATGATGATGATGATGAT 2415
598 TTGTGAGAAATGAGCAAGTAAACCTTGGGAGAGAGAGATGATGATGATGATGATGATGATGATGAT 657
2416 GATCTTGAGCTTACACTTGAATGGAAGGCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 2475

Db 658 GATCTGAGCTGACCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
Qy 2476 CCCAATGCGATCTC 2490
Db 718 CTGTTATGTCATGCTC 732

RESULT 11
LOCUS AY074829 411 bp mRNA linear PLN 04-FEB-2002

DEFINITION Arabidopsis thaliana Atg19350/F18014_4 mRNA, complete cds.

ACCESSION AY074829
VERSION AY074829.1 GI:18491116

KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana (Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 411)

REFERENCE

AUTHORS

Chen, R., Chen, H., Kim, C.J., Meyers, M.C., Ban, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE

Arabidopsis cDNA clones

JOURNAL

REFERENCE

2 (bases 1 to 411)
Chen, R., Chen, H., Kim, C.J., Meyers, M.C., Ban, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Direct Submision
Submitted (14-JAN-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Chen, R., Chen, H., Kim, C.J., Meyers, M.C., Shin, P., Ban, J., Bower, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Chen, R. (SSP/salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/salk) contributed equally to this work as PIs.

FEATURES

source

1..411

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="U12325"

/note="This clone is in pUNI 51"

ecotype: Columbia"

CDS

1..411

/note="unknown protein"
/codon_start=1
/product="AF1919350/F18014_4"
/protein_id="AF169527.1"
/db_xref="GI:18491117"
/translation="MSMAKOSMTSLNTPYAVASAPASPTNHRPHAPATTPECDESD
SSTVDSHMSIFOKAQOQPFSSAVPTSPFNVLKPAPOQLSPNTAIIQETIGSSEF
KFENSOVPEWGERIHDVAMEDELTLTGKXHS"

BASE COUNT 108 a 104 c 89 g 110 t

ORIGIN

Query Match 12.1%; Score 326.4; DB 8; Length 411;
Best Local Similarity 88.5%; Pred. No. 1.5e-58;
Matches 354; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 2058 TGCTAAACATCAATAGGCGCTCTTTAATATCTCTTATGCGGTTTGCACTGCTG 2117
DB 12 TGTAAACAGTCAAGACTCTTTGAACCTACCCGTTTATGCGGTCTGCACTGCCAG 71

QY 2118 TCCGACATCGCCACAGTTTCATACCTGCTCTATACCTGAATGATGAGTCTGA 2177
DB 72 TCTTCTATATATCCCGATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 131

QY 2178 CTCTTCACTGTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2237
DB 132 CTCTTCACTGTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191

QY 2238 ATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2297
DB 192 ATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 251

QY 2298 GCAGATGCTCCAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2357
DB 252 GCATTTGCTCCAAACAGCAGCAATCCAGAGATGCTGCTGCTGCTGCTGCTGCTG 311

QY 2358 TGAGATAGCCAAAGTTAAACCTGCGGAGAGAGAGATACATGATGCTGCTGCTG 2417
DB 312 TGAGAGAGCCAAAGTTAAAGCCATGCGGAGAGAGATCCATGATGCTGCTGCTG 371

QY 2418 TCTTGAAGCTTACATTTGGAATGGAAGCTGCTGCTGCTGCTGCTGCTGCTG 2457
DB 372 TCTTGAAGCTTACATTTGGAATGGAAGCTGCTGCTGCTGCTGCTGCTGCTG 411

RESULT 12
AF395901 1494 bp mRNA linear PLN 08-JUL-2001
LOCUS Lycopersicon esculentum mature anther-specific protein LAT61 mRNA,
DEFINITION complete cds.
ACCESSION AF395901
VERSION AF395901.1 GI:14626760
KEYWORDS
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eubaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
Asteridae; Lemnids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE McCormick, S.
TITLE Tomato mature anther specific protein (LELAT61).
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1494)
McCormick, S.
DIRECT SUBMISSION
SUBMITTED (27-JUN-2001) Plant Gene Expression Center,
USDA/ARS/UC-Berkeley, 800 Buchanan St., Albany, CA 94710, USA
FEATURES
SOURCE location/Qualifiers
1..1494
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
21..1022
/note="similar to protein encoded by GenBank accession

number AC025808"
/codon_start=1
/product="mature anther-specific protein LAT61"
/protein_id="AAK71662.1"
/db_xref="GI:14626761"
/translation="MMWEGSPASSRAGAGSGAGVGIPESGGGGGGRKRSW
RRNNRRRRRAVAKITGTAQGNVNLPRKDNNEVKAICTEGKIVPEPCT
TYRGGCKTPHEIGSTINITPSSRRHSPPSSYRASPIPSIOPSTSSPSPSAD
AMHSHPSFLQNVPSLPRLISNSNAPVTPPLSPTRHPKOTNLTAKESMFAL
NIPIFASAPASPTNVRPPTIPECDESDSTIDSGQINFORVAVSVPSPFN
VKPYPOPRLPMDITDKGKSIDPFENYSVAMEGERIHDVFDLETLTGSGNARI"

BASE COUNT 428 a 329 c 302 g 435 t

ORIGIN

Query Match 7.2%; Score 194.8; DB 8; Length 1494;
Best Local Similarity 60.9%; Pred. No. 1.2e-30;
Matches 454; Conservative 0; Mismatches 232; Indels 60; Gaps 6;

QY 1712 AGGATGCAAGCCTTTACCTGATGATAGCTGGACCTTCATCTCGAGTAACTCCATATT 1771
DB 337 AGGATGCAAGCCTTTACCTGATGATAGCTGGACCTTCATCTCGAGTAACTCCATATT 396

QY 1772 CATCAGAAACAGAGCCCTTTTCATGAGCCTTCAAGTCCATCCATCTTCAAG 1831
DB 397 CTTACGCGCATCAAGTCCCATCATCATCTTGTCTAGCCCAATTCATCTTATCAGC 456

QY 1832 TTAGCCGCTTTCTTATCATTTCCGAGCTCTTCTGCGGTGAACCAATTAACATGT 1891
DB 457 CAAGTCCAACTTCTTCTTCCGAGCTCTTCTGCGGTGAACCAATTAACATGT 510

QY 1892 CCTTACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1951
DB 511 TATCATATCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 567

QY 1952 GAATCTCAACAGTTGTCAGTTACCCCAAGCTCTGAGCCGCTTCAAGAACCGCA 2011
DB 568 GAATCTCAACAGTTGTCAGTTACCCCAAGCTCTGAGCCGCTTCAAGAACCGCA 627

QY 2012 AACGGTCCCTTACCTGGAATCTATCGTAAAGCAATCCATGCGCATTCCTAAACATCA 2071
DB 628 AACCTT-----CAATTGAACATTTGGCAAAAGATCA 663

QY 2072 TGGCGCTTTTATATATCTTTCTATAGGCTTTTGCACTGCTGATGCTGACATCGCC 2131
DB 664 TGTTCCTTTTAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 720

QY 2132 ACAGTTTCATACCTGCTCTTACCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2191
DB 721 TTTACGCTTTTATCTCTTCAATCTTACCCGAGTGAATCTGACTCATTTTACCATTTG 780

QY 2192 ATTCTGCTATGATTAAGCTTTTCAAGATTGCAACAAACAGCATTTCTGCTCTTA 2251
DB 781 ATTACGCGCATGATCAATTTCAAAAGATGCG-----TCA 819

QY 2252 TGGGCGCAACCTCTCTTACCTTCAATCTTGTGAACCTGCGGCTTACAGATGTCTCAA 2311
DB 820 ATGTTTCACTTCTTCAACATTTATCTTGTGAACCTGCGGCTTACAGATGTCTCAA 877

QY 2312 ATATGCTGCTTCTTCAAGATGATGTCAAAGCTCTGAGTTTAAATTGAGATACCAAG 2371
DB 878 -TATGATATGATACAGACAGAGGTAGACATAGACTTTGACTTTAAATGTATCAG 936

QY 2372 TTAACCTTGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2431
DB 937 TCAAGGATGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 996

QY 2432 TTGAATAGGAGAGCTTCTGCTTGA 2457
DB 997 TTGAATAGGAGAGCTTCTGCTTGA 1022

RESULT 13
ATAP22

```

LOCUS       ATAP22               198354 bp    DNA    linear    PLN 07-APR-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, ESSA I AP2 contig fragment
ACCESSION   299708
VERSION     299708.1
KEYWORDS    GI:4006885
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
AUTHORS     Bevan M., Terry N., Voe P., Holm J., L., Mewes H.W., Mayer K.F.X.
            and Schellier C.
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 198354)
AUTHORS     EU Arabidopsis sequencing project.
TITLE       Direct Substitution
JOURNAL     Submitted (07-APR-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schueller@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de,Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@brc.ac.uk
COMMENT     On Dec 12, 1998 this sequence version replaced gi:2464894.
            The annotation of this entry was produced with considerable
            contributions from Stephane Rombauts and Pierre Rouze, Department
            of Genetics, University of Ghent, Ledeganckstraat 35, 9000 Ghent,
            BE, E-mail: stromengengp.rug.ac.be, p.rouze@genp.rug.ac.be
            Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATAP21 at the 5' end.
FEATURES             source
     source          1..198354
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     source          1..10046
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="cosmid C1C7A10.4E12"
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="cosmid C1C7A10.5C09"
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="cosmid C1C7A10.2H02"
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="cosmid C1C7A10.5E02"
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="cosmid C1C7A10.2F03"
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
     source          8316..99612
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="cosmid C1C7A10.5E02"
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="cosmid C1C7A10.2F03"
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
     source          99088..121580
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
     source          108937..198353
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="BAC TBMU10C14"
     gene          8165..10851
                     /gene="C7A10.500"
                     /join(8165..8344,8417..8522,8638..8983,9116..9209,
                     9502..9621,9720..9932,10027..10089,10166..10284,
                     10372..10430,10523..10851)
                     /gene="C7A10.500"
     CDS          8165..10851
                     /note="similar to hypothetical protein T517.13 from
                     Arabidopsis thaliana chromosome II BAC T517 genomic
                     sequence, PATX:G2642165
                     contains Neutral zinc metalloproteinase, zinc-binding
                     region signature [ILAHBMHAW]
                     contains EST gb:N65197, H76120, AA067496"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB16816.1"
                     /db_xref="GI:4006886"
                     /db_xref="SPTREMBL:O23197"
                     /translation="MEFLFLPGITVFLVAGKGLLPMGWLTILKSGSHKSPGQCN
                     GYREDRNIEGRYSNAGSDPKETIECAIALISDEEHVTPDDKCKTIEYKSETE
                     EDDDDDEDEETMRKQLEAEEERKRVAKQIEEEKRAEQLEETKULARLE
                     EEMERKSKQLEDEDLAALQESNVAISPPRYDPNSNITLPVPFLPSHRICGCOA
                     EIGHGRFLAWVAFLGILNVAMHPIPTNAGLIYVAHPMPMKYCSHSDGTPRC
                     SCERMPKDTKYLIDLDGRKLCLECDASIMDIECOPLYLEIREPEGJMKVQESN
                     SYAFGEISIKRAMEGEEKGHNLPETRGICSEBOTVTVLRPRIGACVAKLDMIT
                     EPCRLIRREYVRIILLYGLPRILTSILAHEMHMLRLNGVPLNLRPEEGICQVL
                     AHWLESSTVAGSLVDIASSSSSAVVSSSKGSGSPFKKLGEPFKHOLEBSSSA
                     YDGFQGNQAVLKGLRRLDHLRLTLTTPPKVI"
     exon          8165..8344
                     /gene="C7A10.500"
                     /number=1
     intron        8345..8416
                     /gene="C7A10.500"
                     /number=1
     exon          8417..8522
                     /gene="C7A10.500"
                     /number=2
     intron        8523..8637
                     /gene="C7A10.500"
                     /number=2
     exon          8638..8983
                     /gene="C7A10.500"
                     /number=3
     intron        8984..9115
                     /gene="C7A10.500"
                     /number=3
     exon          9116..9209
                     /gene="C7A10.500"
                     /number=4
     intron        9210..9501
                     /gene="C7A10.500"
                     /number=4
     exon          9502..9621
                     /gene="C7A10.500"
                     /number=5
     intron        9622..9719
                     /gene="C7A10.500"
                     /number=5
     exon          9720..9932
                     /gene="C7A10.500"
                     /number=6
     intron        9933..10026
                     /gene="C7A10.500"
                     /number=6
     exon          10027..10089
                     /gene="C7A10.500"
                     /number=7

```

intron 10090..10165
/gene="C7A10.500"
/number=7
exon 10166..10284
/gene="C7A10.500"
/number=8
intron 10285..10371
/gene="C7A10.500"
/number=8
exon 10372..10430
/gene="C7A10.500"
/number=9
intron 10431..10522
/gene="C7A10.500"
/number=9
exon 10523..10851
/gene="C7A10.500"
/number=10
gene 13598..16004
/gene="C7A10.510"
join(13598..13824,13914..13956,14045..14071,14245..14343,
14416..14528,14855..14990,15133..15234,15322..15417,
15513..15533,15643..15792,15894..16004)
/gene="C7A10.510"
/note="intron number 9 is a special U12 intron
similarity to various predicted proteins of Arabidopsis
and yeast
Contains Somatotropin, prolactin and related hormones
signature (CVLDFILQYIYK)
contains EST gb:H7J118, AA605507"
/codon_start=1
/product="putative protein"
/protein_id="CAB16817.1"
/db_xref="GI:4006887"
/db_xref="SPTRMBL:023198"
/translation="MVSIGYCLKEKTCVWVEIYFDCLCNLDVSPALGINSILC
MGVAEIQVITNFKSSNGVSLFPLAAMDFNVLGSLPATLPQYALTY
VSTVLIQIYDIYIKLCRHRRTKICQDEBDEKRPKPTMGSAISIGGSYK
DSRRFEVTSASRLASGTPRLTSYFRVAKSGLRIMOSTMVRHTRKCRRAF
GTPLAASLPLQAKSLAEKKAHNSRRLNERLVERLSAUGQWGLMAAIYWGGRIP
QIMNIRKSGVEGNPLMFIPLVANAIVGSLVIRTEEDNIKPNLPLDLAIVCV
LDLFIQYIYKCRKISLESREDAVGVESKTFVS"
exon 13598..13824
/gene="C7A10.510"
/number=1
intron 13825..13913
/gene="C7A10.510"
/number=1
exon 13914..13956
/gene="C7A10.510"
/number=2
intron 13957..14044
/gene="C7A10.510"
/number=2

Query Match 5.8; Score 156; DB 8; Length 198354;
Best Local Similarity 51.4; Pred. No. 2,1e-22;
Matches 735; Conservative 0; Mismatches 565; Indels 130; Gaps 11;

QY 1119 CTACGTCGACATCAGCAGCTGCGCGCGCGCGAGCAGCGCGAGGAGGAGCCGT 1178
DB 35200 CTTCAATGCGCGCGCGAGGAGGAGGAGGAGGATCATCTCGGAGCTACTCCGA 35259
QY 1179 CGTGGAGAGAGGAGATTAATCGAGAGAGAGAGAGAGAGAGAGCTGTACTCGGA 1238
DB 35260 CGTGGAG 35319
QY 1239 AGATATACCTGCGCTAGAGCTCAAGTATTAATTTGCTTAACATTGTGATATA 1298
DB 35320 AGATTTACTCTGCTTAGAGCTCAAGGTAACTATAAGCTTCTTAACACTGCGATACA 35379
QY 1299 ATGAAGTCTTAAAGCTCTTGTGTGAAGCTGGTGTGAAGAGATGTACTA 1358

DB 35380 ACAGGTTCTTAAAGCTCTGTCTCGAAGCTGGTGGATCGTGAGAGAGATGACCA 35439
QY 1359 CTTATGCGAGGTAA-----GACTTTCATTTTTCAGATCTGAGCTT 1405
DB 35440 CTTATGCGAGGTAAAGCTGAAATTCGGAATTTCAATTTTAGATCGAATCGAGCTT 35499
QY 1406 -----GTTTATGATGTTTGTGATGTTGAATCTGAATCGTGATTT 1449
DB 35500 CTCCTAATCTCAATTAATCAATGATGATCTGTGCAATGTTGTTTAATTCATTC 35559
QY 1450 CAATTGTGTTAAT-----GGTTGATCTGAAATTTGAGGTTTCTCA 1497
DB 35560 GAATTTAGAGACTGTAGATCTTGAGATTTGATTTCACTTTCATCTGTCATG 35619
QY 1498 AAGTGAATTTGATCATCAGAAACTATGATGATCTGATTTCTCAAGTGAATTTATG 1557
DB 35620 GTTGAATTTCAATTTCCCTATTTGATGATGATCTGATCTGATGATTAAGT 35679
QY 1558 GTT-----TCTTCTAATTTAGGTTATTTAGTATGCTAAGTCTTAATC 1606
DB 35680 GATTATGAATTCAGATGTTTCTGTTTGAATCATCTAATTTATAGATCATGATTTG 35739
QY 1607 TTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1666
DB 35740 ATTATGAGAAATTTTGGGAAATCTTTTGTCTTGAATCTTGAATCAATTAATGAC 35799
QY 1667 TGATTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1702
DB 35800 TTTGACGCTGATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 35859
QY 1703 -----TGAGTATCAGGATGACAGGCTTTACCTGAGATGATGATGATGATGAT 1757
DB 35860 GATTCTTGAGAAACAGGGGTTTACCCACAGATCATGATTTTACAGAAATCTTACAA 35919
QY 1758 AGTACCTCATATTCATCAGAAACAGAGCCCTCTTTCATCAGCTTTCAAGTCCAT 1817
DB 35920 CTTACAGACAAATTCATCATATCCAAAGTCCATCATCTTTCCAGATCTTCC 35979
QY 1818 CCATCTTACCAAGTTAGCCGCTCTTCTTATCATATCCGAGTCTTTCGCGGTGAAC 1877
DB 35980 ACCTTCGTACACAGAAAGTCACTCATCATCTTCCGAGTCACTCGCTATGACGG 36039
QY 1878 AAA--TAACAATATGCTCTATATTC--TTCCCTTCTCAGAAATGATGATCTTCT 1934
DB 36040 AAACCTTCTTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 36099
QY 1935 TTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1994
DB 36100 TAACTTCCACTCTTGAATATCCAAAGTGGGCTGTGATCTTCTTCTTCTTCTTCT 36159
QY 1995 GACTTCTAAGAACCGAAACCGTTCCTTAACTGGAATCTATCGCTAAGCAATCATG 2054
DB 36160 TACTTCTCGGTTCGAGGAGGAACTTCTTCAAGCAATTAACAA----- 36207
QY 2055 CATTCTAACAATCAATGCGCTTTTATTTATCTTCTTCTTCTTCTTCTTCTTCTTCT 2114
DB 36208 -----TGCGGATCTTCTTCAACGTTTGAACATCTGCTTCTTCTTCTTCTTCT 36261
QY 2115 TAGTCGACACATCGCACACAGTTTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2174
DB 36262 TA-----GATCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 36317
QY 2175 ---TGACTTTCAGT 2231
DB 36318 CGAGAGAGATTCATGAGAGATTCAGGAGGTGATCAATTTCAATCTACTGCTCTAC 36377
QY 2232 ACAGCATTTCTGTCTGT 2291
DB 36378 TTCACCAACATTTAACTTGTTCAGCAAACTTCTTATGCGCATTTGATAT----- 36425
QY 2292 GCCTCAGAGATGTCTCAAAATCTGCTGCTTCAAGAGATTTGTCAAGGCTCTGAGTT 2351
DB 36426 -----GAAAGGTCAAGCTGGGAAATGTCTCGAATGATGAGAGAGGAGGAGGAG 36476

QY 2352 TAAATTGAGATAGCGAAGTTAAACCTCGGAGAGAGAGATACATCATGCGGTAT 2411
DB 36477 TGAAGTTGAGAAATGAAACAGTTAGCCGCGGAGAGCGAAATATTCATCAAGTTGGCCG 36536
QY 2412 GGAGAGCTTGAGCTTACACTTGGAATGGAGAGCTCGGTTGACATA 2461
DB 36537 AGAAGATTCGAACTCACTTCGCGGCACTTAAGCAGATGCTGAAAAA 36586

RESULT 14
ATCHR1v86/c 198780 bp DNA linear PLN 16-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 86.
VERSION AL161590 GI:7270623
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 198780)
EU Arabidopsis sequencing project.
Direct Submision
Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
leomech@mp.biochem.mpg.de, mayer@mips.biochem.mpg.de, Proj.
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>
this fragment has an overlap with ATCHR1v85 at the 5' end and an
overlap with ATCHR1v87 at the 3' end.
Location/Qualifiers
1. 198780
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variate="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
12203..13776
/gene="AT4g36750"
join(12203..12795,13548..13776)
/gene="AT4g36750"
/note="strong similarity to minor allergen, Alternaria
alternata, PIR2:84311
contains EST gb:R64949, AA651052"
/codon_start=1
/product="minor allergen"
/protein_id="CA880342.1"
/db_xref="GI:7270624"
/translation="MGKGGCVSPKKKKPAITDGPICIDDDNATNAPIQIDDDQTTI
DGGRTATNTGTTTAAKISPLIFVVFYSMGVSLARMKKGVDSVGG
VEATLVRPETLSEVVEOMKAPVKSLEPEITAAELTADGFLGPPTRVGCMAAM
KAPFDSTGSLMKESLAGKRPAGFPVSTGTGGGQETPAWTAITOLVHGMFLPPIGVT
PGAGMRKMDISIRGSPYGAIVPAGDSREARETELTALAEHQNTMAIVRLAQP"
12203..12795
/gene="AT4g36750"
/number=1
12796..13547
/gene="AT4g36750"
/number=1
13548..13776
/gene="AT4g36750"
/number=2
14283..17541
/gene="AT4g36760"
join(14283..14342,14482..14545,14645..14784,14874..15158,
15237..15632,15717..15803,15889..15953,16062..16128,
16205..16339,16440..16520,16618..16672,16768..16872,
16902..17073,17162..17286,17374..17541)
/gene="AT4g36760"
/note="similarity to cytoplasmic aminopeptidase P, Rattus
norvegicus, PATX:G2760920
contains EST gb:N96008, A1998300.1"
/codon_start=1
/product="aminopeptidase-like protein"
/protein_id="CA880342.1"
/db_xref="GI:7270625"
/translation="MASHSPPLDALVVPSEEDYHQSEVYSARDKRRREFVSGSGAGLA
LITTKEARLMTDGRYFLQALQQLSDMTLMRGEDPLVEWMSDNLPEEANIQVDSNC
VAVDPANRWGSKPAKNOKLITTTDLVDEWVSRPSESPVVAHPAILITTSAPFLYVVK
PFDLRARLKKQKARGLVIALDEVAWLVNIRGTDVAVCPVVAHPAILITTSAPFLYVVK
KVSDEANSTYFNGSLGYEVRBYTDVSDVALLADRLISSASTTYQHEAKKMDIADSD
OPDRLVNDPASCCYALYSKLDKLVLPSPISLAKLPVELEGKLAHVRDAAV
VQYLVLWDNQMOELVXGASGYFLEAENSKKRPSTSKLTVTVADKLESLASKEHFRG
LSFPITSSVGSNAAVIHVSPPEACAMDPDKIYILDSGAYDGLDITDITRVYFGAP
SAHEKCTYAVFPKGVALGNARPKGTNGTYLIDILARAFLMKYGLDYRHGTGCVGSY
LCVHEBQHVSFRPSARNVPLQATWTVTDEPGYEGEDNGFGRILEENLVVNDAREFNF
GDKGTIQFPHITNAPVQVLIIDDELTRSEIDWLNTHSKCKDILAFPMNQTENMLK
KATEPVSA"
14283..14342
/gene="AT4g36760"
/number=1
14343..14481
/gene="AT4g36760"
/number=1
14482..14545
/gene="AT4g36760"
/number=2
14546..14644
/gene="AT4g36760"
/number=2
14645..14784
/gene="AT4g36760"
/number=3
14785..14873
/gene="AT4g36760"
/number=3
14874..15158
/gene="AT4g36760"
/number=4
15159..15236
/gene="AT4g36760"
/number=4
15237..15632
/gene="AT4g36760"
/number=4
15633..15716
/gene="AT4g36760"
/number=5
15717..15803
/gene="AT4g36760"
/number=6
15804..15888
/gene="AT4g36760"
/number=6
15889..15953
/gene="AT4g36760"
/number=7
15954..16061
/gene="AT4g36760"
/number=7
16062..16128
/gene="AT4g36760"
/number=8
16129..16204
/gene="AT4g36760"
/number=8
16205..16339
/gene="AT4g36760"
/number=9
16340..16439

17002..17073,17162..17286,17374..17541)
/gene="AT4g36760"
/note="similarity to cytoplasmic aminopeptidase P, Rattus
norvegicus, PATX:G2760920
contains EST gb:N96008, A1998300.1"
/codon_start=1
/product="aminopeptidase-like protein"
/protein_id="CA880342.1"
/db_xref="GI:7270625"
/translation="MASHSPPLDALVVPSEEDYHQSEVYSARDKRRREFVSGSGAGLA
LITTKEARLMTDGRYFLQALQQLSDMTLMRGEDPLVEWMSDNLPEEANIQVDSNC
VAVDPANRWGSKPAKNOKLITTTDLVDEWVSRPSESPVVAHPAILITTSAPFLYVVK
PFDLRARLKKQKARGLVIALDEVAWLVNIRGTDVAVCPVVAHPAILITTSAPFLYVVK
KVSDEANSTYFNGSLGYEVRBYTDVSDVALLADRLISSASTTYQHEAKKMDIADSD
OPDRLVNDPASCCYALYSKLDKLVLPSPISLAKLPVELEGKLAHVRDAAV
VQYLVLWDNQMOELVXGASGYFLEAENSKKRPSTSKLTVTVADKLESLASKEHFRG
LSFPITSSVGSNAAVIHVSPPEACAMDPDKIYILDSGAYDGLDITDITRVYFGAP
SAHEKCTYAVFPKGVALGNARPKGTNGTYLIDILARAFLMKYGLDYRHGTGCVGSY
LCVHEBQHVSFRPSARNVPLQATWTVTDEPGYEGEDNGFGRILEENLVVNDAREFNF
GDKGTIQFPHITNAPVQVLIIDDELTRSEIDWLNTHSKCKDILAFPMNQTENMLK
KATEPVSA"
14283..14342
/gene="AT4g36760"
/number=1
14343..14481
/gene="AT4g36760"
/number=1
14482..14545
/gene="AT4g36760"
/number=2
14546..14644
/gene="AT4g36760"
/number=2
14645..14784
/gene="AT4g36760"
/number=3
14785..14873
/gene="AT4g36760"
/number=3
14874..15158
/gene="AT4g36760"
/number=4
15159..15236
/gene="AT4g36760"
/number=4
15237..15632
/gene="AT4g36760"
/number=4
15633..15716
/gene="AT4g36760"
/number=5
15717..15803
/gene="AT4g36760"
/number=6
15804..15888
/gene="AT4g36760"
/number=6
15889..15953
/gene="AT4g36760"
/number=7
15954..16061
/gene="AT4g36760"
/number=7
16062..16128
/gene="AT4g36760"
/number=8
16129..16204
/gene="AT4g36760"
/number=8
16205..16339
/gene="AT4g36760"
/number=9
16340..16439

[illegible]

Db	21778	CTTCAATGCGCGCCGGAGAGAGAGAGAGAGATCAATCGTGGGACGTACTCCGA	21719
Qy	1179	CGTGGAGAGAAAGGAGAAATATATCGAGAGAGAGAAACGGAGAAAGACTGTAGCTGGCA	1238
Db	21718	CGTGGAAAGAGAGAGAAACAATAAGAGAGAAAGAAAGAAAGAGAGCCACTCTGCTA	21658S
Qy	1239	AGATATACATGGGCTTAGAGCTCAAGGTGATTAAATTTGGCCTTAACATGTGTAAATA	1298
Db	21658	AGATTACTGTGGCTTAGAGCTCAAGGTAACTTAAGCTTCTTAAGACATCGCGATTACA	21599S
Qy	1299	ATGAGAGCCTTAAAGCCTTTGTGTAAAGTGCTTTGGGTGTGGAAGAAGTGTACTA	1358
Db	21598	ACAGAGTCTTAAAGCTCTGTCTGAAGCTGTGTGATCGTGGAAACAGATGCGACCA	21539S
Qy	1359	CTTATCGCAAGGTGAA-----GACTTCTCAATTTTTCAGATCTGAGCTT	1405
Db	21538	CTTATCGCAAGGTAAACTGAAATTCGGATTTCAATTTTAACTCGGAATCTGAGCTT	21479S
Qy	1406	-----GTTTATGATGTTTGTAGTGTGMAATCTGAATCGTATTTT	1449
Db	21478	CTCTCAATTACTATTAATCAATCGATGATCTGTGGATGTGTGTTTAAATTCATTC	21419S
Qy	1450	CAATGTGTTAAT-----GGTTTGAATCTGAAATTTGAGGTTTTCCTCA	1497
Db	21418	GAATTTTAAAGACTCTTGAGATCTTGGATTTGGATTTTCACTTTTCACTGTTGATTGG	21359S
Qy	1498	AGGTGATTTGAATCATCAAGAACTAAGATGATGATTCGATTTCTCAAGTAAATTTATGG	1557
Db	21358	GTTGAATTTTCAATTTCCCTCTATTGATGATGAATCTGCTTCTGATGATTAAGAACT	21289S
Qy	1558	GTT-----TCTTTCTAATTTTAAAGTTATTAATGATGCTAAAGCTTTAATC	1606
Db	21298	GATTATGAATTCGATGATTTCTCTGTTGATCATCTAATTCATATGATCATGTAAGTTTG	21239S
Qy	1607	TTTTATGATGATCTGTGTCCAAAGTCATGCTATGTGTTCTTTTGTCTTACCTGTGAT	1666
Db	21238	ATTATTAAGAAATTTTGGGAAATCTCTTTTGTCTTTGATGCTTGAATAAATGATGAC	21179S
Qy	1667	TGATTGATGTTGATGTTATGTTTGTGTTGT-----	1702
Db	21178	TTTGAAGGTTGATCATCATGTGTGTTTATTAATGATCTAATCTTTGCTCACTAAATGTT	21119S
Qy	1703	-----TGAATATCAGGAGATGCAAGCTTTTACCTGTGATGATGCTGGACCTTCATCTCG	1757
Db	21118	GATTTCTGTGGAAACAGGGGTTTAGCCACAGATCAGATTAATTCAGAACTCCTACAAA	21059S
Qy	1758	AGTAACTCATATTTATCATCAGAAACAGAGCCTCTTTATCAGCCTTTAAGTCCCAT	1817
Db	21058	CTTCAGCAAAATATCATCATCCMACCAAGTCCACATCATCAGCTTTTCCAAAGTCTGCG	20899S
Qy	1818	CCCATCTTCAAGATGAGCCGCTCTTTCATCATTTCCGAGTCTTCTGCGGTGAACC	1877
Db	20898	ACCTTCGTACAGAGAGTCCAGTCTCATCATCTTCCGAGTCCATCTGCTATGACGG	20939S
Qy	1878	AAA--TAACACATGTCTCTACATTC--TTCCCTTCTCAGAAATGTTGGAATTCCTTC	1934
Db	20938	AAACCTTCTTATACCTTCTTCTTCTGCTCCAGACAAACATGTGCTTCTGATTCCTGC	20879S
Qy	1935	TTCTCTTCTTCCCTCAGAAATCTCAACAGTGTCCAGTTAACCCACCGGTCTCATGCGC	1994
Db	20878	TAACTTCACTCTTGAATATCCAAAGTGGCGCTGTGACCTCTCCGTTGTCATCTCC	20819S
Qy	1995	GACTTCTAAGAACCGAAACCGTGGCTAACTGGGAATCTATGCGTAAAGCAATCATGGC	2054
Db	20818	TACTTCTGTGTGTAAGCGGAAACTTACAGAGCAATTAACAAA-----	20771S
Qy	2055	CATTGCTAAACATATGCGCTCTTAAATATCTTTCTATGCGGTTTCTGACCTGC	2114
Db	20770	-----TGGCGGCTCTTAAACAGTTTGAACATCGCTTTGCGATCTCTGCTCATC	20717S
Qy	2115	TAGTCCGACATGCGCACAGTTTCAATCCTGTGCTACTAATCTGAATGTGATAGTC	2174
Db	20716	TA---GACTGTCGTGTGTCAACCAACGCACTTCAATAACCGGAATGTGATAGTC	20661S

QY 2175 ---TGACTCTTCCACTGTGTTATCTGTCATTGGATAGCTTTCAGAAAGTTGCACACAA 2231
DB 20660 CGAGGAGAGATTCGATCGAGGATTCAGGAAAGGTGATCAATTCATCTACTGCTCTAC 20601
QY 2232 ACAGCCATTCCTCGCTCTATGAGTGGCAACCTCTCTACTTCACTTGTGAAACCTGC 2291
DB 20600 TTACCCACACATTAACTTGTTCAGCAAACTTCTATGSCCATGTAT------ 20553
QY 2292 GCCTCAGCAGATGTCTCCAAATGCTGCTCTCCAGAGATGTGCTCAAGCTCTGACTT 2351
DB 20552 -----GAAAGGTGAGCTGGGGAATGTCGGATGATGAGGAGAGGCGAGACTT 20502
QY 2352 TAAATTGAGAAATGCCAAGTTAAACCTGGAGAGAGAGATCACTGATGCGGTAT 2411
DB 20501 TGACTTGAGAAATGAGACGATTAAGCCCTGGGAAGCGCAAAATGATCATGAAGTTGGCGT 20442
QY 2412 CGAGGATCTTACGCTTACCTTGAATGGGAAAGGCTGTGGTACATA 2461
DB 20441 AGAAGATCTCGAATCTCTGCGTGCGACTTAAGCCAGATGCTGAAAAA 20392

RESULT 15
ATT3A5/c 84196 bp DNA linear PLN 26-JAN-2000
LOCUS Arabidopsis thaliana DNA chromosome 3, BAC clone T3A5.
DEFINITION AL132979
ACCESSION AL132979
VERSION AL132979.2 GI:6782244
KEYWORDS
SOURCE Arabidopsis thaliana (chale cress).
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Bleecker, H., Mewes, H.W., Lemcke, K., Mayer, K.F.X., Queciet, F. and Salanoubat, M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 84196)
AUTHORS EU Arabidopsis sequencing project.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mpi.biochem.mpg.de, mayer@mpi.biochem.mpg.de Project Coordinator: Marcel Salanoubat and Francis Queciet, Groupe ment d'Interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue Gaëton Cremieux, BP191, 91006 Evry Cedex, France; http://www.genoscope.cns.fr
COMMENT On Jan 27, 2000 this sequence version replaced gi:6434247. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES
Source 1. 84196
Location/Qualifiers
1. 84196
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
2983..3338
/note="355 bp LTR"
3339..7564
/note="LTR-terminoson"
7565..7922
/note="367bp LTR"
12339..13197
/gene="T3A5.10"
join(12339..12665,12742..12803,12888..13055,13125..13197)
/gene="T3A5.10"
/note="similarity to cyclin-dependent kinase inhibitor (ICK1), Arabidopsis thaliana, EMBL:ATU94772_1"
/codon_start=1
/product="putative protein"

protein_id="CAB62432.1"
/db_xref="GI:6561966"
/translation="MAVRRRRRVNVEGVTTTTRKRKKEEVVDLVESRIILSPCV
QATNRGISTVARNAGASETSVIVYRRDSPPVEQCOIEEDSSVSCSTSEBSKRR
IEFVDLEENNDDREIETSMIVDILNKSEBSMNDSVAVEDVESRRRLKSLHETV
KEALEBDFPQVAEADILRNKLLCSMKYNDFEKDEPIGCGRYEWKLPN"
12339..12665
/gene="T3A5.10"
/number=1
12666..12741
/gene="T3A5.10"
/number=1
12742..12803
/gene="T3A5.10"
/number=2
12804..12887
/gene="T3A5.10"
/number=2
12888..13055
/gene="T3A5.10"
/number=3
13056..13124
/gene="T3A5.10"
/number=3
13125..13197
/gene="T3A5.10"
/number=4
13198..16470
/gene="T3A5.20"
complement(join(15858..15901,16014..16470))
/gene="T3A5.20"
complement(join(15858..15901,16014..16470))
/gene="T3A5.20"
/note="similarity to predicted proteins, Arabidopsis thaliana"
/codon_start=1
/product="putative protein"
/protein_id="CAB62433.1"
/db_xref="GI:6561967"
/translation="MARLEMINSNKGAARISFNEFVIRSEKSNKSNNTNRSRFS
MPSADRFASVDSMTIPADEIPLKGLILPFKTSVHRTGLGELLTEEGSVNDGTF
SLRPLTSSSFSTGTRRELLGLKRTVRSKTKDKNVEVLSDHKIISGNVATREC
QVADR"
complement(15858..15901)
/gene="T3A5.20"
/number=1
complement(15902..16013)
/gene="T3A5.20"
/number=1
complement(16014..16470)
/gene="T3A5.20"
/number=2
18303..19931
/gene="T3A5.30"
complement(18303..19931)
/gene="T3A5.30"
complement(18303..19931)
/gene="T3A5.30"
/codon_start=1
/product="acarecrow-like 7 (SCL7)"
/protein_id="CAB62434.1"
/db_xref="GI:6561968"
/translation="MAVWCTDSGNLMAIAOQIKKQKQOOSQHQOEOEOEPNPNP
BSFGFTLPDSGSDPFOVNTDGFPHLEHHQNAVAEEDSDMESLJINGDS
OTNPDPPIYGHDPFVSPFRLASPSYLNKVNQDASQQLPPPASTAISSPPSPQ
HPPPPPODPDNLNPIFKAIHDAKPKTKPTDLIRIKESVSESDPIORVGYFPE
ALSHKTEPSSSSSSLEDPILISYTKLADACYSEFAULTAOALILEATNQNHHI
VDFGIFQCIQMSALLOALTRSSGKPTRIRIGIPAPSLGDSRPSLILATGNLRDA
ALIDLNFEFYPVLTPIQLNIGSSFRVDPDEVLVNVMLELYKLIDETATVGAALRA
RLNLRIVTLGAEVSLNVEFANRVNLSRFSAVESLEPVLDRDSKERLVERVL
FGRRLINDVSDDNKPKTRFGLMKEKQWRVLMKAGFEVPKPSNYAVSAOKLLLM
NYSYTLVLSVESEPGFISLANNVPLLTVSSWR"
complement(18303..19931)

exon

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 19:11:09 ; Search time 683 Seconds
(without alignments)
10619.898 Million cell updates/sec

Title: US-09-995-938A-2

Perfect score: 2687
Sequence: 1 attcagtttgcctcaattc.....attcatttcagataaaaaa 2687

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1068.2	39.8	1501	21	AAC37477
2	532.4	19.8	1359	21	AAC37137
3	532.4	19.8	1418	21	AAC36495
4	149	5.5	831	24	ABZ12580
5	141.6	5.3	642	21	AAC42317
6	121.4	4.3	413	25	ABX22943
7	110.8	4.1	1213	21	AAC51043
8	110.8	4.1	1217	21	AAC42550

9	109.6	4.1	284	25	ABX32572
10	93.2	3.5	459	22	AAH87878
11	70	2.6	6145	24	ABL32972
12	69.6	2.6	5499	24	ABQ65971
13	65	2.4	47108	24	ABK31510
14	62.6	2.3	6343	25	ABZ10194
15	61.6	2.3	6343	25	ABZ10194
16	60.8	2.3	16033	24	ABL33404
17	60.6	2.3	5678	24	ABL33138
18	60.4	2.2	6070	24	ABQ67130
19	60.4	2.2	6070	24	ABL70372
20	60.4	2.2	6070	24	ABL33679
21	60.4	2.2	6070	24	ABL34579
22	60.4	2.2	6954	24	ABL33391
23	60.4	2.2	14006	24	ABL33958
24	60.2	2.2	5430	22	AA546292
25	60.2	2.2	6283	24	ABL32088
26	60.2	2.2	113515	24	ABL34174
27	60	2.2	612	22	AAH71471
28	60	2.2	15373	24	ABL32465
29	59.6	2.2	5449	22	AA546384
30	59.6	2.2	5449	24	ABK40008
31	59.6	2.2	5449	24	ABL32849
32	59.6	2.2	16891	24	ABNB0059
33	59.6	2.2	16891	24	ABL32653
34	59.6	2.2	16891	24	ABL34505
35	59.4	2.2	529	25	ABX52180
36	59.4	2.2	20486	24	ABL34611
37	59.4	2.2	34769	22	AA546774
38	59.2	2.2	580	24	ABQ34864
39	59.2	2.2	580	24	ABQ34865
40	59.2	2.2	6112	24	ABL32489
41	59	2.2	446	25	ABX38235
42	59	2.2	5195	24	ABL32921
43	59	2.2	6351	24	ABL33747
44	59	2.2	6944	24	ABK34027
45	59	2.2	9539	22	AA545347

ALIGNMENTS

RESULT 1
AAC37477
ID AAC37477 standard; DNA; 1501 BP.
XX
AC AAC37477;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17521.
XX
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.

Human GDP-mannose
Peppermint plant o
Human immune syste
Human angio genesis
Signal transductio
Human immune syste
Haematopoietic cel
Human immune syste
Human immune syste
Human angio genesis
Chemically treated
Human immune syste
Human metastasis a
Human immune syste
Human immune syste
Tumour suppressor
Human immune syste
Human cervical can
Human immune syste
Tumour suppressor
Human chemically p
Human immune syste
Human chemically m
Human immune syste
Human metastasis a
Bovine EST associat
Human metastasis a
Tumour suppressor
Oligonucleotide fo
Human immune syste
Bovine EST associat
Human immune syste
Human immune syste
Human DNA for stag
Chemically pretrea

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140053.
PR 23-JUN-1999; 99US-0140054.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151203.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158399.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 39.8%; Score 1068.2; DB 21; Length 1501;
Best Local Similarity 80.4%; Pred. No. 2.1e-248;
Matches 1424; Conservative 0; Mismatches 3; Indels 344; Gaps 1;
QY 917 ATTGTGGAGAGAGAGAGATCTCTTCTTCATTCAGAGAGAGAGAGAGATTT 976
DB 1 ATTGTGGAGAGAGAGAGATCTCTTCTTCATTCAGAGAGAGAGAGAGATTT 60
QY 977 CCTGTGAGAGAGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1036
DB 61 CCTGTGAGAGAGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 1037 AATCTTTACCCCACTTATCAAAAGCTCTCTCCGAGATTTCGAGGGTGGTGGT 1096
DB 121 AATCTTTACCCCACTTATCAAAAGCTCTCTCCGAGATTTCGAGGGTGGTGGT 180
QY 1097 TTCCCGATGACTTCGATGAGAGAGAGATTCAGATCAGACAGCTGAGCTGGGGGGG 1156
DB 181 TTCCCGATGACTTCGATGAGAGAGAGATTCAGATCAGACAGCTGAGCTGGGGGGG 240
QY 1157 GCAGCGGCGAG 1216
DB 241 GCAGCGGCGAG 300
QY 1217 CGAG 1276
DB 301 CGAG 360
QY 1277 TTGCTTAAATTTGATTAATTAAGAGCTCTTAAAGCTCTTGTGTTGAAGCTGG 1336
DB 361 TTGCTTAAATTTGATTAATTAAGAGCTCTTAAAGCTCTTGTGTTGAAGCTGG 420
QY 1337 GTTGTGAAG 1396
DB 421 GTTGTGAAG 451
QY 1397 TCTGAGCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1456
DB 452 TCTGAGCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 451
QY 1457 GGTAAATGGGTTGAATTCAGAAATTTGAGGGTTTCTCAAAGTGAATTTGAATCA 1516
DB 452 GGTAAATGGGTTGAATTCAGAAATTTGAGGGTTTCTCAAAGTGAATTTGAATCA 451
QY 1517 GAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1576
DB 452 GAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 451

QY 1577 AGTATATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1636
DB 452 AGTATATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 451
QY 1637 TGCATTTGTTCTTTTGTCTTACCTGTATGATGATGATGATGATGATGATG 1696
DB 452 TGCATTTGTTCTTTTGTCTTACCTGTATGATGATGATGATGATGATGATG 451
QY 1697 TTTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1756
DB 452 TTTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 496
QY 1757 GAGTAATCCATATTTCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1816
DB 497 GAGTAATCCATATTTCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556
QY 1817 TCCCATCTTACAG 1876
DB 557 TCCCATCTTACAG 616
QY 1877 CAAATTAACAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1936
DB 617 CAAATTAACAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 676
QY 1937 CTCTTCTTCCCTCAGAGATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1996
DB 677 CTCTTCTTCCCTCAGAGATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736
QY 1997 CTCTTAAGAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2056
DB 737 CTCTTAAGAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 796
QY 2057 TTGCTTAAACAAG 2116
DB 797 TTGCTTAAACAAG 856
QY 2117 GTCCGACATGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2176
DB 857 GTCCGACATGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 916
QY 2177 ACTCTTCACTGTTGATTTGATGATGATGATGATGATGATGATGATGATGAT 2236
DB 917 ACTCTTCACTGTTGATTTGATGATGATGATGATGATGATGATGATGATGAT 976
QY 2237 CATTTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2296
DB 977 CATTTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1036
QY 2297 AGGAGATGCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2356
DB 1037 AGGAGATGCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1096
QY 2357 TTGAGATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2416
DB 1097 TTGAGATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1156
QY 2417 ATCTTGAAG 2476
DB 1157 ATCTTGAAG 1216
QY 2477 CCAAAATGAG 2536
DB 1217 CCAAAATGAG 1276
QY 2537 CAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2596
DB 1277 CAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1336
QY 2597 CCAAGCTTCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2656
DB 1337 CCAAGCTTCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1396

QY 2657 GTTCAAGCATTTGTGATTAATAAAA 2687
Db 1397 GTTCAAGCATTTGTGATTAATAAAA 1427

RESULT 2
AAC37137
ID AAC37137 standard; DNA; 1359 BP.

AC AAC37137;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16310.

XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 18-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137500.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139859.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144335.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145182.

PR 23-JUL-1999; 99US-0145185.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

```

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149422.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157863.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 19.8%; Score 532.4; DB 21; Length 1359;
 Best Local Similarity 18.8%; Pred. No. 1.2e-118;
 Matches 640; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

```

QY 1712 AGGATGCAAGCCTTACTGCTGATGAGATGCTGGAGCTTCATCTGAGTAACCTCATATT 1771
DB 374 AGGACACAAAGCCTTACTGCTGATGAGATGCTTCGAGCACTCTTACT 433
QY 1772 CATCAGAACCAAGCCTTCTTTCATGAGCTTTCGAAGTCCATCCATCTTCCAG 1831
DB 434 CTTCCTATTAACCAAGCCTTCTTTCATGAGTCCATCTTATCTTACCAAG 493

```

```

QY 1832 TTAGCCGCTCTCTTCATCATCCCGAGTCCCTTCGCGGGAACCAATTAACAATAT 1891
DB 494 TCAGTCTCTCTCTCTTCATTCGCCAGTCTTTCAGTTGGATCCACAAATAT-- 551
QY 1892 CCTTCATATTCCTCCCTTCCTTCCTTCAGAAATGGTGGCATCTCTTCTTCTTCTCTCA 1951
DB 552 -CTCACAATCTTCCTCTTCCTTCAGAAATGGTGGATTCCTTCATTCGCTTCTTCATTA 610
QY 1952 GAATCTAAACAGTTCCTTCAGTTACCCACCGGCTCATCGCCGACTTCTAAGAACCCGA 2011
DB 611 GAATCTCAACAGTCTTCCTTCAGTCCACAGTGTATCCCACTTCTAAGAACCTCA 670
QY 2012 AACGCTGCTTACCTGGGAATCTATGGCTAAGCAATCATGGCCAT--TGTAAACAAT 2068
DB 671 AACCATTCCTTACCTGGGAATCTTTCACAAACATCATATGCTTCATGCTGTCAACAGT 730
QY 2069 CAATGGCTCTTTCATATTCCTTCCTTCATGCGGTTTCTGCACTGCTAATGCGACATC 2128
DB 731 CAATGACTTCTTTCATATTCCTTCCTTCATGCGGTTTATGCGGTTGCGACCTGCTTCTATC 790
QY 2129 GCCACAGTTTCATATTCCTTCCTTCATGCGGTTTCTGCACTGCTAATGCGACATC 2188
DB 791 ATCCGCAATTCATGCTCCGCTTACTATATCTGATGATGATGATGATGATGATGATGATG 850
QY 2189 TTGATTCGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2248
DB 851 TTGATTCGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 910
QY 2249 CTATGTCGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2308
DB 911 CTATGTCGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 970
QY 2309 CAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2368
DB 971 CAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1030
QY 2369 AAGTTAAACCTGGGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 2428
DB 1031 AAGTTAAACCTGGGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 1090
QY 2429 CAATTCGAAATGGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2488
DB 1091 CGCTTGGAAATGGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
QY 2489 TC 2490
DB 1151 TC 1152

```

RESULT 3
 AAC36495
 ID AAC36495 standard; DNA; 1418 BP.

AC AAC36495;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14021.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134216.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 26-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147326.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

```

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 19.8%; Score 532.4; DB 21; Length 1418;

Best Local Similarity 81.8%; Pred. NO. 1.2e-118; Matches 640; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

```

QY 1712 AGGAGTCAAGACCTTTACCTGTGAGATAGCTGGGACTTCTGATCTGATCAATTCATATT 1771
DB 449 AGGAGTCAAGACCTTTACCTGTGAGATAGCTGGGACTTCTGATCTGATCAATTCATATT 508
QY 1772 CATCAGAAACGAGGACCTTCTTTCATCAGCCTTTCAAATGCCATCCATCTTACCAAG 1831
DB 509 CTTCACATTAACCAAGCTCTTCTTCTTCCACTTTTGTAGAGCCCATCTTATCTTACCAAG 568
QY 1832 TTAGCCGCTTTTTCATCATTCGCCGCTCTTCCGGGGAACCAATTAACAAGAT 1891
DB 569 TCAGTCTTCTTCTTCTTCTTCAATCCGAGCTCTTCTGAGTGTGATCCACACATAT 626
QY 1892 CCTCTACATCTTCCCTTCTTCTGAGAAATGGGACTTCTTCTTCTTCTTCTTCTTCTTCA 1951
DB 627 CTTCACATCTTCTTCTTCTTCTGAGAAATGGGACTTCTTCTTCTTCTTCTTCTTCTTCA 685
QY 1952 GAATCTCAACAGTGTTCAGTTACCCACCGGCTTCATGCGGACTTCTTCAAGACCGCA 2011
DB 686 GAATCTCAACAGTGTTCAGTTACCCACCGGCTTCATGCGGACTTCTTCAAGACCGCA 745
QY 2012 AACCGTTCCTTACTGGGAATCTATCGCTAAGCAATCATGGCCAT--TGCCTAAACAT 2068
DB 746 AACCGTTCCTTACTGGGAATCTTTCACAAACATCATGCTTCAATGCTGCTAAACAT 805
QY 2069 CAATGGGCTTTTAAATTAATATCTTCAATGCGGCTTCTGACCTGCTGCTGCTGCTGCTG 2128
DB 806 CAATGGGCTTTTAAATTAATATCTTCAATGCGGCTTCTGACCTGCTGCTGCTGCTGCTG 865
QY 2129 GCCACAGATTTCATACCTGCTACTATACCTGAATGATGATGATGATGATGATGATGATG 2188
DB 866 ATGCGCAGTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 925
QY 2189 TTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2248
DB 926 TTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 985

```

```

QY 2249 CTATGTCGCAACCTCTCTTACCTTCAATCTTGTAAACCTGCGCTCAGAGATGCTC 2308
DB 986 CTATGTCGCAACCTCTCTTACCTTCAATCTTGTAAACCTGCGCTCAGAGATGCTC 1045
QY 2309 CAATATGCTGCTCTTCCAGAGATTTGTCAAAAGCTCTGAGTTAAATTTGAGAAATGCC 2368
DB 1046 CAATATGCTGCTCTTCCAGAGATTTGTCAAAAGCTCTGAGTTAAATTTGAGAAATGCC 1105
QY 2369 AAGTTAAACCTTGGAGAGAGAGATTCATGATGATGATGATGATGATGATGATGATGATG 2428
DB 1106 AAGTTAAACCTTGGAGAGAGAGATTCATGATGATGATGATGATGATGATGATGATGATG 1165
QY 2429 CAATGGAATGGAAGGCTCGGTTCATTAATTAACCACTAGGCAACCAATGCGATG 2488
DB 1166 CAATGGAATGGAAGGCTCGGTTCATTAATTAACCACTAGGCAACCAATGCGATG 1225
QY 2489 TC 2490
DB 1226 TC 1227

```

RESULT 4

ABZ12580 standard; DNA; 831 BP.

ABZ12580;

21-JAN-2003 (first entry)

Arabidopsis thaliana stress regulated gene SEQ ID NO 385.

Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

Arabidopsis thaliana.

WO200216655-A2.

28-FEB-2002.

24-AUG-2001; 2001WO-US26685.

24-AUG-2000; 2000US-227866P.

26-JAN-2001; 2001US-264647P.

22-JUN-2001; 2001US-300111P.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA PARTICIPATIONS AG.

Harper JF, Kreps J, Wang X, Zhu T;

WPI; 2002-304127/34.

Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses

Claim 144; SEQ ID NO 385; 577bp + Sequence Listing; English.

The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising:

(a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and

(b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stresses. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ1574) used in methods of the invention.

Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office.

PR	18-MAY-1999	99US-0134768
PR	19-MAY-1999	99US-0134769
PR	20-MAY-1999	99US-0134911
PR	21-MAY-1999	99US-0135154
PR	24-MAY-1999	99US-0135353
PR	25-MAY-1999	99US-0135639
PR	27-MAY-1999	99US-0136021
PR	28-MAY-1999	99US-0136282
PR	01-JUN-1999	99US-0136782
PR	03-JUN-1999	99US-0137222
PR	04-JUN-1999	99US-0137558
PR	07-JUN-1999	99US-0137502
PR	08-JUN-1999	99US-0137724
PR	10-JUN-1999	99US-0138094
PR	10-JUN-1999	99US-0138085
PR	14-JUN-1999	99US-0138847
PR	16-JUN-1999	99US-0139119
PR	18-JUN-1999	99US-0139452
PR	18-JUN-1999	99US-0139453
PR	18-JUN-1999	99US-0139457
PR	18-JUN-1999	99US-0139458
PR	18-JUN-1999	99US-0139459
PR	18-JUN-1999	99US-0139460
PR	18-JUN-1999	99US-0139461
PR	18-JUN-1999	99US-0139462
PR	18-JUN-1999	99US-0139463
PR	18-JUN-1999	99US-0139750
PR	18-JUN-1999	99US-0139763
PR	21-JUN-1999	99US-0139817
PR	22-JUN-1999	99US-0139899
PR	22-JUN-1999	99US-0140053
PR	23-JUN-1999	99US-0140054
PR	24-JUN-1999	99US-0140255
PR	26-JUN-1999	99US-0140823
PR	29-JUN-1999	99US-0140991
PR	30-JUN-1999	99US-0141887
PR	01-JUL-1999	99US-0141254
PR	01-JUL-1999	99US-0141254
PR	02-JUL-1999	99US-0142055
PR	06-JUL-1999	99US-0142303
PR	08-JUL-1999	99US-0142803
PR	09-JUL-1999	99US-0142927
PR	12-JUL-1999	99US-0143242
PR	13-JUL-1999	99US-0143362
PR	14-JUL-1999	99US-0143424
PR	15-JUL-1999	99US-0144005
PR	16-JUL-1999	99US-0144085
PR	19-JUL-1999	99US-0144331
PR	19-JUL-1999	99US-0144332
PR	19-JUL-1999	99US-0144334
PR	19-JUL-1999	99US-0144352
PR	20-JUL-1999	99US-0144352
PR	20-JUL-1999	99US-0144632
PR	20-JUL-1999	99US-0144684
PR	21-JUL-1999	99US-0144614
PR	21-JUL-1999	99US-0145086
PR	21-JUL-1999	99US-0145086
PR	22-JUL-1999	99US-0145087
PR	22-JUL-1999	99US-0145087
PR	22-JUL-1999	99US-0145089
PR	22-JUL-1999	99US-0145145
PR	23-JUL-1999	99US-0145145
PR	23-JUL-1999	99US-0145145
PR	26-JUL-1999	99US-0145276
PR	27-JUL-1999	99US-0145313
PR	27-JUL-1999	99US-0145313

```

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```

```

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5 3%; Score 141.6; UB 21; Length 642;
Best Local Similarity 72.0%; Pred. No. 3,9e+24;
Matches 183; Conservative 1; Mismatches 70; Indels 0; Gaps 0;

QY 1119 CTACGTCGACATCAGCAGCTGCAGCGCGCCGACAGCGCGAGAGAGACCGT 1178
DB 128 CTTMATGCGCCCGGAGAGAGAGAGAGAGAGATCATCGTGGAGCATCTCGA 187
QY 1179 CGTGAGAGAAAGGGAATATCGAGAGAGAAAGACCGAGAAAGCTGTACCTCGA 1238
DB 188 CGTGAAAGAGAGAGAAATTAAGAGAGAGAAAGAAAGAGAGCCATCACTGCTA 247
QY 1239 AGATATACACTGGCTTAGAGCTCAAGTCATTAATTGCCATAACATTGTGATATA 1298
DB 248 AGATTACTCTGCTTTAAGCTCAAGTCACTAATTAAGCTTCTTACGACTGCGATACA 307
QY 1299 ATGAGTCTTAAAGCTTTGTGTAAGCTGTGGTGTGTAAGAAAGATGTAATA 1358
DB 308 ACGAGTCTTAAAGCTCTCTGTCGAAGCTGTGGATCGTGAAGAGATGCAACA 367
QY 1359 CTTATCGCAAGGTG 1372
DB 368 CTTATCGCAAGGGG 381

RESULT 6
ABX22943
ID ABX22943 standard; cDNA; 413 BP.
AC ABX22943;
XX 10-FEB-2003 (first entry)
XX Human GDP-mannose 4,6-dehydratase (GM4, 6D) DNA #5000.
DE Human GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW complex carbohydrate; gene replacement therapy; immunosuppressive;
KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW antiaesthetic; vasotropic.
XX Homo sapiens.
XX OS
XX US2002110548-A1.
XX
XX 15-AUG-2002.
XX
XX 11-JUN-2001; 2001US-0878574.
XX
XX 22-NOV-1996; 96US-0753233.
XX 03-DEC-1997; 97US-0984246.
XX 09-SEP-1998; 98US-0149674.
XX 14-JUN-1999; 99US-0333177.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Sullivan F, Kitz R, Kumar R;
XX
XX WPI; 2003-066673/06.
XX
XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
XX peptide, for manufacturing complex carbohydrates, or as targets for
XX screening GM4,6D antagonists for treating e.g. arthritis, or transplant
XX rejection
PT

```

PR 01-APR-1999; 99US-0127462

PR	06-APR-1999	99US-01288234
PR	08-APR-1999	99US-01288714
PR	16-APR-1999	99US-01298455
PR	19-APR-1999	99US-01300477
PR	21-APR-1999	99US-01304494
PR	23-APR-1999	99US-01305010
PR	23-APR-1999	99US-01308511
PR	28-APR-1999	99US-01311449
PR	30-APR-1999	99US-01312408
PR	30-APR-1999	99US-01324078
PR	04-MAY-1999	99US-01324454
PR	05-MAY-1999	99US-01324645
PR	06-MAY-1999	99US-01324666
PR	06-MAY-1999	99US-01324677
PR	07-MAY-1999	99US-01328853
PR	11-MAY-1999	99US-01341256
PR	14-MAY-1999	99US-01342169
PR	14-MAY-1999	99US-01344219
PR	14-MAY-1999	99US-01344231
PR	14-MAY-1999	99US-01343700
PR	18-MAY-1999	99US-01347968
PR	19-MAY-1999	99US-01349411
PR	20-MAY-1999	99US-01351244
PR	21-MAY-1999	99US-01353533
PR	24-MAY-1999	99US-01356629
PR	27-MAY-1999	99US-01360211
PR	27-MAY-1999	99US-01363621
PR	28-MAY-1999	99US-01367812
PR	01-JUN-1999	99US-01372782
PR	01-JUN-1999	99US-01376288
PR	03-JUN-1999	99US-01375578
PR	04-JUN-1999	99US-01375002
PR	07-JUN-1999	99US-01377744
PR	08-JUN-1999	99US-01380094
PR	10-JUN-1999	99US-01385400
PR	10-JUN-1999	99US-01388647
PR	11-JUN-1999	99US-01393191
PR	16-JUN-1999	99US-01394612
PR	17-JUN-1999	99US-01394633
PR	17-JUN-1999	99US-01394932
PR	18-JUN-1999	99US-01394945
PR	18-JUN-1999	99US-01394955
PR	18-JUN-1999	99US-01394960
PR	18-JUN-1999	99US-01394961
PR	18-JUN-1999	99US-01394962
PR	18-JUN-1999	99US-01394963
PR	18-JUN-1999	99US-01394964
PR	18-JUN-1999	99US-01394965
PR	18-JUN-1999	99US-01394966
PR	18-JUN-1999	99US-01394967
PR	18-JUN-1999	99US-01394968
PR	18-JUN-1999	99US-01394969
PR	18-JUN-1999	99US-01394970
PR	18-JUN-1999	99US-01394971
PR	18-JUN-1999	99US-01394972
PR	18-JUN-1999	99US-01394973
PR	18-JUN-1999	99US-01394974
PR	18-JUN-1999	99US-01394975
PR	18-JUN-1999	99US-01394976
PR	18-JUN-1999	99US-01394977
PR	18-JUN-1999	99US-01394978
PR	18-JUN-1999	99US-01394979
PR	18-JUN-1999	99US-01394980
PR	18-JUN-1999	99US-01394981
PR	18-JUN-1999	99US-01394982
PR	18-JUN-1999	99US-01394983
PR	18-JUN-1999	99US-01394984
PR	18-JUN-1999	99US-01394985
PR	18-JUN-1999	99US-01394986
PR	18-JUN-1999	99US-01394987
PR	18-JUN-1999	99US-01394988
PR	18-JUN-1999	99US-01394989
PR	18-JUN-1999	99US-01394990
PR	18-JUN-1999	99US-01394991
PR	18-JUN-1999	99US-01394992
PR	18-JUN-1999	99US-01394993
PR	18-JUN-1999	99US-01394994
PR	18-JUN-1999	99US-01394995
PR	18-JUN-1999	99US-01394996
PR	18-JUN-1999	99US-01394997
PR	18-JUN-1999	99US-01394998
PR	18-JUN-1999	99US-01394999
PR	19-JUN-1999	99US-01444331

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130610.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134841.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135153.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0144055.
PR 06-JUL-1999; 99US-0144390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 110.8; DB 21; Length 1217;
Best Local Similarity 70.5%; Pred. No. 1.5e-16;
Matches 148; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1161 CGCGAGAGAGAAAGCCGCTGAGAGAGAAAGGAAATATCGAGAGAGAGAAAGCGGA 1220
DB 155 CGCGAGATTAAGACCCGACTTGGAAGAGAGAAACAAACGCGAGAGCGCGAA 214
QY 1221 GAAGAGCTGAGCTGCGAAGATATACACTGGGCTTAGAGCTCAAGGTGATATATTGC 1280
DB 215 GAGAGGAGATTGCGGCTAAGATCTTCGAGAGACTAAGATTCATGAAACTTCAAGCTCC 274
QY 1281 CTAAACATTGATATATATGAAGTCTTAAAGCTTTGTTGAAGCTGGTGGTTG 1340
DB 275 CTAAACACTGCGAACAATGATAGTCTTCAAGCTTTATGCAATGATGAGCTGGTGGACTG 334
QY 1341 TTGAAGAAGATGTAATCTTATTCGCAAG 1370
DB 335 TAGAAGACGACGAACTTATTCGCAAG 364

RESULT 9

ABX32572
ID ABX32572 standard; cDNA; 284 BP.

XX ABX32572;

DT 11-FEB-2003 (first entry)

DB Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #14629.

XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; se; inflammation;
KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW complex carbohydrate; gene replacement therapy; immunosuppressive;
KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW antiaesthetic; vasotrophic.
XX Homo sapiens.
OS
XX

PN US2002110548-A1.
XX 15-AUG-2002.
PD 11-JUN-2001; 2001US-0878574.
XX 22-NOV-1996; 96US-0753233.
PR 03-DEC-1997; 97US-0984246.
PR 09-SEP-1998; 98US-0149674.
PR 14-JUN-1999; 99US-0331177.
XX (GENY) GENETICS INST INC.
XX Sullivan F, Kriz R, Kumar R;
PI WPI; 2003-066673/06.
DR

PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT peptide, for manufacturing complex carbohydrates, or as targets for
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT rejection

XX Disclosure; SEQ ID NO 14631; 6pp; English.

XX The invention relates to a composition comprising a human GDP-mannose
XX 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
XX GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
XX in a mammalian subject and for treating or ameliorating diseases affected
XX by the level of cellular fucosylation or diseases affected by the
XX fucosylation of glycoconjugates. These diseases include arthritis,
XX transplant rejection, asthma, sepsis, reperfusion injury, stroke or
XX infection. The GM4,6D peptide or a polynucleotide encoding it is also
XX useful for manufacturing complex carbohydrates and as targets for
XX screening small molecule antagonists of the activity of the enzyme. The
XX polynucleotide is useful in developing an assay for defects in the
XX enzyme, as well as in gene replacement therapy. Sequences
XX ABX17942-ABX17944 and ABX17947-ABX17948 represent DNA molecules encoding
XX human GM4,6D peptides of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html.

XX Sequence 284 BP; 95 A; 62 C; 89 G; 38 T; 0 other;

Query Match 4.1%; Score 109.6; DB 25; Length 284;
Best Local Similarity 73.3%; Pred. No. 1.5e-16;
Matches 154; Conservative 0; Mismatches 54; Indels 2; Gaps 1;

QY 1163 GCGAGAGAGAAAGCCGCTGAGAGAGAAAGGAAATATCGG-AGGAGAGAAAGCGGA 1220
DB 42 GCGAGAGAAAGCCGCTGAGAGAGAAAGGAAACCAAGAAAGAGAAAGAGCGAA 101
QY 1221 GAAGAGCTGAGCTGCGAAGATATACACTGGGCTTAGAGCTCAAGGTGATATATTGC 1280
DB 102 GAAGAGCAATGAGCAGCGAAATATATCTCAGACTTCAAGCTCAGGGAACCTTCAACTTGC 161
QY 1281 CTAAACATTGATATATATGAAGTCTTAAAGCTTTGTTGAAGCTGGTGGTTG 1340
DB 162 CAAAGCATTTGAGAACCAACGAGCTTTGAAAACTTCTGCGAGAAAGCTGGTGGTGG 221
QY 1341 TTGAAGAAGATGTAATCTTATTCGCAAG 1370
DB 222 TAGAAGAGAGGAGACCACTTATTCGCAAG 251

RESULT 10

AAH87878
ID AAH87878 standard; cDNA; 459 BP.

XX AAH87878;

AC 25-SEP-2001 (first entry)

XX
DT
XX

DE Human immune system associated gene SEQ ID NO: 945.

XX Human; immune system disease; cytosine methylation; antiaesthetic;
KW antiarteriosclerotic; antinaemic; cyostatic; nocrotic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antiporiatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene, ds.

XX Homo sapiens.

OS WO200200928-A2.

PN 03-JAN-2002.

PD 02-JUL-2001; 2001WO-EP07537.

PP 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

PA Olek A, Piepenbrock C, Berlin K;

PL WPI; 2002-130909/17.

DR Nucleic acid comprising fragment of chemically modified gene, useful
XX for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation.

XX Claim 1; SEQ ID NO 945; 32bp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

XX SQ Sequence 6145 BP; 1493 A; 99 C; 1309 G; 3244 T; 0 other;

QY Query Match 2.6%; Score 70; DB 24; Length 6145;
Best Local Similarity 50.9%; Pred. No. 2.4e-06;
Matches 166; Conservative 0; Mismatches 160; Indels 0; Gaps 0

Db 1378 TTTCTCCATTTTTCAGATCTGAGCTGTGTTATATGATGTTTGTGATGTTGAATCTGA 1433
Y 5420 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5476

Y 1438 ATTCGTGATTTCAATTCGTGTTAAATGGTTGAACTCGAGATTTGAGGTTTTCGA 1497
Db 5480 TTTTGTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 5533

Y 1498 AAGTGAATTTGAATCATCAGAACTATGATGATCGATCTTCACAAAGTAATTTATG 1553
Db 5540 TATTCGTTTTTTTTTTTATTTTATTTTATTTTATTCGTTTTTATTTTATTTTATTTT 5595

Y 1558 GTTTTCTTTCTAATTTTAGAGTATTTATGGATGCTAAAGCTTAACTTTATGATAG 1611
Db 5600 TTTTATTTATTTATTTATTTATTTATTTATTTATTTTTCGTTTTTTTTTTTTTTTT 5658

Y 1618 ATACTGTGTCACAAAGCATTCGATGTTGTTCTTTGCTTACCGTATGATGATGTT 1677
Db 5660 TTAATTTTATTTATTTTATTTATTTATTTTATTTTATTTTATTTTATTTATTTATG 5719

Y 1678 TGAATGGTTATTTGTTTGTGTTTGT 1703
Db 5720 TTGCTTTTTTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5745

CC sequences of the invention are useful for the diagnosis and therapy of
CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK3159-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX Sequence 47108 BP; 13283 A; 429 C; 9740 G; 23656 T; 0 other;

Query Match 2.4%; Score 65; DB 24; Length 47108;
Best Local Similarity 49.7%; Pred. No. 0.0001;
Matches 222; Conservative 0; Mismatches 220; Indels 5; Gaps 2;

QY 1258 AGCTCAGAGTATATATATGCTTAACTGTGATATATGAAAGTCTTAAAGCTCT 1317
DB 195 AGGTAAAGAAATTTTAAATTTTATATATATATATATATATATATATATATAT 254
QY 1318 TTGTGTTGAAGCTGTGGGTGTGAGA-AGATGCTACTTATTCGCAAGTGAAG 1376
DB 255 ATGTATATATATTTTATTTTATATATATATATATATATATATATATATATAT 314
QY 1377 CTTTCTCATTTCACATCTGAGCTGTGTATATATATATATATATATATAT 1436
DB 315 TTTTGAATTTTGTATTTTATAGTAGTATATATATATATATATATATATATAGT 374
QY 1437 AATCTGTTATTCATATGCTTAAATGGGTTGAATCTGAAATTTGAGGTTTCTC 1496
DB 375 AATAAGTTGTTTATTTTATTTTAAATGGTTATGTTAGTTAGTTATTTT 434
QY 1497 AAGTGAATTTGAATCATCAGAACTATGATGATCTGATTTCTCAAGTGAATTTAG 1556
DB 435 TTGATGATATGAGT---ATTATAGTATATAGTGTATATATATATATATATAT 490
QY 1557 GGTTCCTTCATATTTAGTATATATATATATATATATATATATATATATATAT 1616
DB 491 AATATTTTATATATTTTATGTTTATGTTTATATATATATATATATATATAT 550
QY 1617 GATCTGTTCCAAAGTATGATGTTGTTTCTTTGCTTACCTGATGATGATGAT 1676
DB 551 GTAAATTTGTTAAATTTGTTAGATTTTATATATATATATATATATATATATG 610
QY 1677 TTGATGTTATTTGTTTCTTTTCTT 1703
DB 611 TTTTATATATTTTATGTTTATTTT 637

RESULT 14

ABL32357
ID ABL32357 standard; DNA; 9642 BP.

AC ABL32357;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 330.

XX Human; immune system disease; cytosine methylation; antiaesthetic;
KW antiaesthetic; antianaemic; cytosine; neoplastic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antineumatic; antiaesthetic; antidiabetic; antipsoriatic;
KW antineumatic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

OS Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX

PF 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

PA Olek A, Piepenbrock C, Berlin K;

PI WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation.

PS Claim 1; SEQ ID NO 330; 32pp + Sequence listing; German.

XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

XX Sequence 9642 BP; 2381 A; 88 C; 2349 G; 4824 T; 0 other;

Query Match 2.3%; Score 62.6; DB 24; Length 9642;
Best Local Similarity 49.5%; Pred. No. 0.00018;
Matches 161; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 1386 TTTTTCAGATCTGAGCTGTTTATATATATATATATATATATATATATATAT 1445
DB 3491 TTTTATTTTATTTTAAATTTTAAAGATATATATATATATATATATATATAT 3550
QY 1446 ATTCAATGTTGTTAAATGTTGTTGAATCTGAAATTTGAGGTTTCTCAAGTGAAT 1505
DB 3551 TTTTATTTTATTTTAAATTTTATTTGTTTATTTATTTATTTTATTTTATTT 3610
QY 1506 TTGAATCATCAGAACTATGATGATCTGATTTCTCAAGTGAATTTATGTTTCTT 1565
DB 3611 TAAATTTATTTAGATATATTTTATATTTTCTAGATTTTAAATTTTAAATATTT 3670
QY 1566 TCTAATTTAGATTTATTTATTTGATGCTAAAGCTTTAACTTTTATTTATTT 1625
DB 3671 TTAATTTAGGTTTATTTTATTTATTTATTTTATTTTATTTTATTTTATTT 3730
QY 1626 TCCAAAGTATGATGATGTTTCTTTCTTAACTGTTGATTTGATTTGATTTG 1685
DB 3731 TTTTATTTTATTTTGTTTTAAAGTATATATATATATATTTTATTTTATTT 3790
QY 1686 TATGTTTGTCTTTGTTGAGTAT 1710
DB 3791 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3815

RESULT 15

AB210194/c
ID AB210194 standard; DNA; 6343 BP.

AC AB210194;

DT 16-JAN-2003 (first entry)

DE Haematopoietic cell proliferation disorder related DNA sequence #334.

XX Human; haematopoietic cell proliferation disorder; cytosine;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.

XX Homo sapiens.

XX

PN WO200277272-A2.
 XX 03-OCT-2002.
 PD 26-MAR-2002; 2002WO-EP03401.
 XX 26-MAR-2001; 2001US-278333P.
 PR (EPIC-) EPIGENOMICS AG.
 PA Berlin K, Braun A, Distler J, Guecig D, Howe A, Mueller J,
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E,
 PI Lewin A, Lipsecher E, Mater S, Model F, Mueller V, Otto T,
 PI Pelet C, Schwabe I, Ziebarth H;
 XX WPI: 2003-018942/01.
 DR Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent
 PT that distinguishes between methylated and non-methylated CpG
 PT dinucleotides -
 XX Claim 28; SEQ ID 334; 117bp; English.
 PS
 CC The present invention describes a method for detecting and
 CC differentiating between hematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB209861 to AB21118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used for
 CC differentiating between healthy hematopoietic cells and proliferative
 CC disorder hematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of hematopoietic cell proliferation disorder related
 CC DNA sequences. The nucleotide sequences from the present invention can
 CC also be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC hematopoietic cell proliferative disorders. The present method enables
 CC a highly specific classification of hematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients.
 CC
 XX
 XX Sequence 6343 BP; 1685 A; 0 C; 1623 G; 3035 T; 0 other;
 SQ
 Query Match 2.3%; Score 61.6; DB 25; Length 6343;
 Best Local Similarity 43.1%; Pred. No. 0.00026;
 Matches 295; Conservative 0; Mismatches 389; Indels 0; Gaps 0;
 QY 89 AGACATTAGTACATGCTCTGCACTTCTACAAATAAAGCTCATATCCAAAATATT 148
 DB 3223 ACACCTTAATTAATACAAAATAATATTATTTCCAACTAAATAAATAAATAATTTTCA 3164
 QY 149 ACATGATCATATCATATATGCGCCGCAACCTTGTTAGGACAACCTGTAACCCCT 208
 DB 3163 TTCTTAACCTCCAAAATATATCTCTTTTCCAAATTTAACCACTTACTCTTTAAACAA 3104
 QY 209 TTTTCCTTTTATGTTCAATGAATATACAAAGTTTGGTATGAAATACATAAATAATGATG 268
 DB 3103 TTTTAAATTAACATATCTATTAATAATATTTTCTTTTACTCTTTATATATATATTT 3044
 QY 269 GACCCAGCAATTAATCCAAATTTGGATATAGATTAAGCTTAATAATCAACATGTAA 328
 DB 3043 ATCTTACTTAACATAAATAAATAAATACTTTACATTAATTAATTAATCAAAATTCAT 2984
 QY 329 CCAAACTAAATCTTTATAGACATATGTAATGATATTCACCAATCTTTATATCATTTGT 388
 DB 2983 ACATTTAAATAATTTTATTAATAAATAATTAATTAATTAATAAATAATTTCTTACATTAT 2924

QY 389 AAGTACGAAGAGGTAAAAAAGAGAGCCAGCTGATCATACAACTAATTCAGGACAA 448
 DB 2923 TTTTATCTATCATCAAAAATACTTTATATACACAAAATATCATTTTAAACAAA 2864
 QY 449 AGTAGCAAAATGTTGTTCTTAAAGTGAGATTGTTGATGCAAAAAAGTGATTAATTTTA 508
 DB 2863 ATCCATCTAAATAATCCAAATACCAATATATCAAAAAACAAAACATTAACAAAATAA 2804
 QY 509 TTGAATATATCATATGATGTTAATCAACAGGCTCAGCTATTAATAAATTTTAA 568
 DB 2803 TTTAAATTTTATTTAT 2744
 QY 569 ACCAATCAAAATGTTGTTGTTTTCGTCACAAAGTAAGGACCCACATGAAAGAGCT 628
 DB 2743 AAAAATTAATATACCACTTATTAATTAATTTTACTTAACCAAAAATTAACCT 2684
 QY 629 CCCAGTGTCT 688
 DB 2683 TAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2624
 QY 689 TCACAGAAAAAATAAAGATTAGAACATTAACACAGTTACTAGCGTATTCCTGTG 748
 DB 2623 ACACCTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2564
 QY 749 CACCTTAACATACACCTCTTATAT 772
 DB 2563 AACATATTTTATTTCTTAAAT 2540

Search completed: December 19, 2003, 20:57:01
 Job time : 687 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 20:30:29 ; Search time 5408 Seconds

(without alignments)
12075.837 Million cell updates/sec

Title: US-09-995-938A-2

Perfect score: 2687
Sequence: 1 accttgccttcgacctc.....atcttcgttagataaaaaa 2687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estcin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_pmg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	717.6	26.7	941	B12579	P22H5-T7 IG
2	641.4	23.9	685	AV782667	AV782667
3	473	17.6	639	AV822043	AV822043
4	462	17.2	463	AV440785	AV440785

5	442	16.4	539	9	AV442375
C	416.6	15.5	796	13	B0238047
C	386.2	14.1	1619	9	AV440980
C	378	14.4	1648	9	AV545003
9	356	13.2	1585	9	AV551486
10	352.8	13.1	514	14	T42881
C	323.6	12.0	773	28	BH673471
C	313.8	11.7	568	9	AV782917
C	303.8	11.3	307	14	U74178
C	288.8	10.7	307	9	AA389830
C	280	10.4	602	29	B2476824
C	277.6	10.3	604	10	BE662806
C	277.6	10.3	772	28	B20877
C	270.8	10.1	537	9	A1997341
C	262.8	9.8	643	14	CA934625
C	249.6	9.3	511	28	BH597004
C	227.2	8.5	427	9	AV542332
C	210.4	7.8	511	10	BF277350
C	201.8	7.5	629	13	BQ407744
C	201.6	7.5	541	12	B1136769
C	200.4	7.5	692	28	BH500797
C	196	7.3	456	29	CNSO0ND
C	195.8	7.3	582	14	CA933175
C	194.6	7.2	590	9	AV822219
C	193.8	7.2	801	12	B1433025
C	193.4	7.2	812	14	CB292897
C	193.2	7.2	315	28	BH634176
C	191	7.1	599	10	BG447420
C	188.2	7.0	626	13	BQ407763
C	186.8	6.9	504	12	BH533355
C	185.2	6.9	451	9	AV532273
C	183	6.8	359	28	BH645352
C	181.8	6.8	766	29	BZ476333
C	178.4	6.6	661	28	BH428064
C	173.8	6.5	458	12	B1136953
C	172.4	6.4	400	9	AV734606
C	170.8	6.4	481	9	AV807059
C	166.8	6.2	606	13	BQ591032

ALIGNMENTS

RESULT 1
B12579
LOCUS
DEFINITION
P22H5-T7 IGF Arabidopsis thaliana genomic clone P22H5, genomic survey sequence.
ACCESSION
B12579
VERSION
B12579.1 GI:2093659
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 941)
Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.
BAC End Sequences at ATGC
TITLE
JOURNAL
COMMENT
Other_GSSs: P22H5-Sp6
Contact: Ecker, J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu

Db 505 TCAGAGTTGGCAACACAGCATCTCTGCTCATGCTGCAACCTCTCTACTT 446
 QY 2274 CAATCTTTGAAAACCTGCGCTCAGCAGATGTCCTCAAAATACCTGCTTCCAGAGAT 2333
 Db 445 CAATCTTTGAAAACCTGCGCTCAGCAGATGTCCTCAAAATACCTGCTTCCAGAGAT 366
 QY 2334 TGGTCAAAAGCTCGAGTTTAAATTTGAAATGAGCAAGTAAACCTGGGAGAGAGAG 2393
 Db 385 TGGTCAAAAGCTCGAGTTTAAATTTGAAATGAGCAAGTAAACCTGGGAGAGAGAG 326
 QY 2394 GATACATGATGAGGATGAGAGATCTTGACCTTACCTGAAAATGGAAGGCTGCTG 2453
 Db 325 GATACATGATGAGGATGAGAGATCTTGACCTTACCTGAAAATGGAAGGCTGCTG 266
 QY 2454 TTGACATTAACCACTAGGCAAAACCAATGSCATGTCATTTGAAATATGAGAACTAATCC 2513
 Db 265 TTGACATTAACCACTAGGCAAAACCAATGSCATGTCATTTGAAATATGAGAACTAATCC 206
 QY 2514 TCTTGAGTATTTCTTTCTTGTCCAGATTTTGGATCTTTATGGAATCTCATATGTTCTT 2573
 Db 205 TCTTGAGTATTTCTTTCTTGTCCAGATTTTGGATCTTTATGGAATCTCATATGTTCTT 146
 QY 2574 CACTTATTAATCAAAATATGCTGCGCAAGGCTTCTCATGAG -AAGCATGAGGTTAGA 2632
 Db 145 CACTTATTAATCAAAATATGCTGCGCAAGGCTTCTCATGAGGAGTTGAGTGTAGA 86
 QY 2633 GTGGTATTAATCAATTCATGAATTTGGTTTCAAAAGCATTAATTTGATATTAATAAAAA 2687
 Db 85 GTGGTATTAATCAATTCATGAATTTGGTTTCAAAAGCATTAATTTGATATTAATAAAAA 31

RESULT 3
 AV822043 639 bp mRNA linear EST 01-Apr-2002
 LOCUS AV822043 RAF14 Arabidopsis thaliana cDNA clone RAF14-20-E20 5'
 DEFINITION mRNA sequence.

ACCESSION AV822043
 VERSION AV822043.1 GI:19864075
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 639)
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Saito, M., Nakajima, M., Arai, K., Seki, T., Shimada, K., Shimada, A., Muramatsu, M., Hayashizaki, Y., and Shimozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)

TITLE Unpublished
 JOURNAL Contact: Motoaki Seki
 COMMENT Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: meeki@rc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SacI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES
 source Location/Qualifiers

1..639
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF14-20-E20"
 /dev_stage="rossette plants"
 /lab_host="SOUR"
 /clone_id="RAF14"

BASE COUNT 168 a 144 c 153 g 174 t
 ORIGIN /note="Site 1: Set1; Site 2: Xho1; subjected to cold-created(1,2,5,10,24 hr)"

Query Match 17.6%; Score 473; DB 9; Length 639;
 Best Local Similarity 100.0%; Pred. No. 2.4e-65;
 Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 898 AGATGAAAAAATATTAATTTGTTGAGAGAAAGAGAGATTCTTCTTGATTC 957
 Db 2 AGATGAAAAAATATTAATTTGTTGAGAGAAAGAGAGATTCTTCTTGATTC 61
 QY 958 AGCGAAGAAAACCGATTCCTGAGACCTAATCTTCTCTCTTCTTCTT 1017
 Db 62 AGCGAAGAAAACCGATTCCTGAGACCTAATCTTCTCTCTTCTTCTT 121
 QY 1018 CATCAGCTACGCTTCACAACTTTTCCACCTATTCAAGGCTCTCCGAAATTC 1077
 Db 122 CATCAGCTACGCTTCACAACTTTTCCACCTATTCAAGGCTCTCTCCGAAATTC 181
 QY 1078 GAGGGGTGGTTGGTTTCCGATGACTTCGATGAGCTACGTCATCAGCAGC 1137
 Db 182 GAGGGGTGGTTGGTTTCCGATGACTTCGATGAGCTACGTCATCAGCAGC 241
 QY 1138 TGCAGCTGCGGCGCGGACGACGCGGACGAGAGAACCGCTCGTGAGAGAGAGAA 1197
 Db 242 TGCAGCTGCGGCGCGGACGACGCGGACGAGAGAACCGCTCGTGAGAGAGAGAA 301
 QY 1198 TAATCGAGAGAGAAAGACGAGAAAGAGCTGAGCTGAGAGATATACACTGGGCTTAA 1257
 Db 302 TAATCGAGAGAGAAAGACGAGAAAGAGCTGAGCTGAGAGATATACACTGGGCTTAA 361
 QY 1258 AGCTCAAGGTATTAATTTGCTTAAACATTTGATTAATGAACTCTTAAAGCTCT 1317
 Db 362 AGCTCAAGGTATTAATTTGCTTAAACATTTGATTAATGAACTCTTAAAGCTCT 421
 QY 1318 TTGCTGTAAGCTGTTGGTTGTTGAGAAAGATGTAATCTTATGCAAG 1370
 Db 422 TTGCTGTAAGCTGTTGGTTGTTGAGAAAGATGTAATCTTATGCAAG 474

RESULT 4
 AV440785/c 463 bp mRNA linear EST 14-NOV-2000
 LOCUS AV440785 Arabidopsis thaliana above-ground organ two to six-week
 DEFINITION old Arabidopsis thaliana cDNA clone Ap209b10_E 3', mRNA sequence.
 ACCESSION AV440785
 VERSION AV440785.1 GI:7611157
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 463)
 Asamizu, E., Nakamura, Y., Sato, S., and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries

TITLE DNA Res. 7, 175-180 (2000).
 JOURNAL MEDLINE 20363093
 PUBMED 10907847
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
 source Location/Qualifiers

1..463
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"

/db_xref="taxon:3702"
/clone="AP209b10_f"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_1ib="Arabidopsis thaliana above-ground organ two to six-week old"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI, Site_2: XhoI"

BASE COUNT 139 a 96 c 92 g 135 t 1 others

Query Match 17.2%; Score 462; DB 9; Length 463;
Best Local Similarity 99.8%; Pred. No. 1.4e-63;
Matches 462; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2224 GCACAACACAGCCATCTCTGCTCTATGTCGCAACCTCTCTCACTTCAATCTGTG 2283
DB 463 GCACAACACAGCCATCTCTGCTCTATGTCGCAACCTCTCTCACTTCAATCTGTG 404
QY 2284 AAACCTGCCCTCAGAGATGTCCTCAATATCTGCTGCTTCCAGAGATTGCTCAAGC 2343
DB 403 AAACCTGCCCTCAGAGATGTCCTCAATATCTGCTGCTTCCAGAGATTGCTCAAGC 344
QY 2344 TCTGAGTTAAATTGAGATAGCCAAAGTTAAACCTGGGAGAGAGAGATGATGAT 2403
DB 343 TCTGAGTTAAATTGAGATAGCCAAAGTTAAACCTGGGAGAGAGAGATGATGAT 284
QY 2404 GTGGGATGAGAGATCTTACCTTACCTTGAAGAGAGAGAGAGAGATGATGAT 2463
DB 283 GTGGGATGAGAGATCTTACCTTACCTTGAAGAGAGAGAGAGAGATGATGAT 224
QY 2464 CAATGAGCAAAACCAATGAGATGATGATGATGATGATGATGATGATGATGAT 2523
DB 223 CAATGAGCAAAACCAATGAGATGATGATGATGATGATGATGATGATGATGAT 164
QY 2524 TTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2583
DB 163 TTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104
QY 2584 CCAATATGCTGCCCAAGCCCTTCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 2643
DB 103 CCAATATGCTGCCCAAGCCCTTCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 44
QY 2644 ATTCAATGATTTGTTTCAAAAGCATTTATTTGATGATGATGATGATGATGAT 2686
DB 43 ATTCAATGATTTGTTTCAAAAGCATTTATTTGATGATGATGATGATGATGAT 1

RESULT 5
AV442375 539 bp mRNA linear EST 14-NOV-2000
LOCUS AV442375 Arabidopsis thaliana above-ground organ two to six-week
DEFINITION old Arabidopsis thaliana cDNA clone AP209b10_r 5', mRNA sequence.
ACCESSION AV442375
VERSION AV442375.1 GI:7612783
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 539)
REFERENCE Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
JOURNAL MEDLINE
PUBMED 10937847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
FEATURES
source
1..539
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="AP209b10_f"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_1ib="Arabidopsis thaliana above-ground organ two to six-week old"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI, Site_2: XhoI"

BASE COUNT 141 a 118 c 137 g 143 t

Query Match 16.4%; Score 442; DB 9; Length 539;
Best Local Similarity 100.0%; Pred. No. 2.1e-60;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 929 AGAAGAGAGATTTCTTCTTCTGATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 988
DB 1 AGAAGAGAGATTTCTTCTTCTGATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 989 TAATCTTCACT 1048
DB 61 TAATCTTCACT 120
QY 1049 ACCATTTCAAAGCT 1108
DB 121 ACCATTTCAAAGCT 180
QY 1109 TCGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1168
DB 181 TCGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 1169 AGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1228
DB 241 AGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 1229 GTAGCTGCAAGATATATATATATATATATATATATATATATATATATAT 1288
DB 301 GTAGCTGCAAGATATATATATATATATATATATATATATATATATATAT 360
QY 1289 TGTGATTAATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1348
DB 361 TGTGATTAATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 1349 GATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1370
DB 421 GATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 442

RESULT 6
BU238047/c 796 bp mRNA linear EST 06-SEP-2002
LOCUS BU238047/c
DEFINITION Des01_14f06 A/Ds01 AAFc ECORC cold stressed flaxseed seedlings
ACCESSION BU238047
VERSION BU238047.1 GI:22749872
KEYWORDS EST.
SOURCE Descurainia sophia
ORGANISM Descurainia sophia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Descurainia.
1 (bases 1 to 796)
REFERENCE Singh, J.A., Piche, C., Couroux, P., De Moors, A., Harris, L.J., Hattori,
J.I., Ouellet, T., Robert, L.S., Spott, D. and Tinker, N.A.
Expressed Sequence Tags from Cold-Stressed Descurainia sophia
Seedlings
Unpublished

OY	2196	TGGTATTGGATAAGCTTTCAAAATTTGGCAAAACAGCAATCTCTGCTCTAATGT	225
Db	379	TGGTATTGGATAAGCTTTCAAAAGTTTGCACAAACAGCAATCTCTGCTCTAATGT	320
OY	2256	GCCAACTTCTCCTACCTTCAATCTTGTGAACCTCGCTCAGCAGATGTCTCCAAATAC	2315
Db	319	GCCAACTTCTCCTACCTTCAATCTGTGAACCTCGCACCACAGCAATGTCTCCAAACAC	260
OY	2316	TGCTGCTCTCCAAAGATTGGTCAAAGCTCTGAGTTTAAATTGGAAATAGCCAAATTAA	2375
Db	259	AGCAGCAATCCAAAGAGATTGTCAAAGCTCCAGATTTAAGTTTGAAACAGCAAGTTAA	200
OY	2376	ACCTTGGGAAGGAGAGAGATCATGATGTGGTATGAGAGATCTTGAAGCTTACACTTGG	2435
Db	199	GCCATGGGAAGGAGAGAGATCCATGATGTGCTATGAGGAGATCTAAGAGCTCAGCTTGG	140
OY	2436	AAATGGGAAGGCTCGTGGTTGACATTAACAATTAGGCAAAACCCAAATGGCATGTC	2490
Db	139	AAATGCTAAAGCTCAATAGTGGATGAAGATATACATGAACCTGTTATGTATCATGTC	85

RESULT 8

LOCUS	AV45003	648 bp	mRNA	linear	EST 07-SEP-2000
DEFINITION	AV45003 Arabidopsis thaliana roots Columbia Arabidopsis thaliana				
ACCESSION	CDNA clone R267c11F 3', mRNA sequence.				
VERSION	AV45003				
KEYWORDS	AV45003.1	GI:8716417			
SOURCE	EST.				
ORGANISM	Arabidopsis thaliana (thale cress)				
	Arabidopsis thaliana				
	Arabidopsis thaliana				

REFERENCE	AUTHORS	TITLE
1 (pages 1 to 648)	Aasatun, E., Nakamura, Y., Sato, S. and Tabata, S.	Large scale analysis of cDNA in <i>Arabidopsis thaliana</i> : Generation

JOURNAL
DNA RES. 7, 175-180 (2000)

COMMENT Contact: Erika Asamizu

FEATURES	Location/Qualifiers
SOURCE	1. .648

BASE COUNT	187 a	136 c	151 g	174 t
ORIGIN				

Query Match	14.1%	Score 379;	DB 9;	Length 648;
Best Local Similarity	64.1%	Pred. No. 3e-50;		
Matches 439;	Conservative 0;	Mismatches 80;	Indels 3;	Gaps 1;

Oy 1972 GTTACCCCAACCGGTCTCATCGCCGCACTTAAAGAACCCGAAACCGTTGGCTAACTGGGAA 2031
 Db 648: GTCTCTCCACACAGTGTCTATCCCACTTATGAAACCCCAACCATTTGGCTACTTGGGAA 589

Oy 2032 TCTATCGCTAAGCAATCCATGGGCAT---TGTAAACAATCATATGGCGCTTTTAATAT 2088
 Db 588 TCTTTTACCAACATCCATGTCATCGCTGTCTAAACAGTCATGACTCTCTTGAACATAC 529

QY	2089	CCITTTCTATGCGGGTTCTTGTGCACTGTCTAGTCCGACACATGCCACCAAGTTTCATACCCCTG	2148
Db	528	CCGTTTATATCGGGTGTGTGCACTGCGCAGTCCATACATCATGCGCAGTTCCATGCTCCG	469
QY	2149	GCTACTATACCTGATATGTATGTAGTCTGTGACTCTTCCACTGTGATTTCTGGTCAATTGGATA	2208
Db	468	GCTACTATACCTGAATGTATGTAGTCTGTGACTCTTCCACTGTGATTTCTGGTCAATTGGATA	409
QY	2209	AGCTTTCAGAAAGTTTGGACACAAACACCACTTCTCTGCTCTATGTGGTCCCACTCTCTCT	2268
Db	408	AGCTTTCAAAAGTTTGGACACAAACACCACTTCTCTGCTCTATGTGGTCCCACTCGCCT	349
QY	2269	ACCTTCAATTTGTGAAACCTGGGGCTCAGCACAATGTCTTCCAAATACTGCTGCTTCCAA	2328
Db	348	ACCTTCAATTTGTGAAACCTGGACACACAGCAATTTGTCTTCCAAACACGACGCAATTCCAA	289
QY	2329	GAGATTTGTCNAAGCTCTGAGTTTAATTGAGATATAGCCAACTTTAAACCTGGGGAAGA	2388
Db	288	GAGATTTGTGTAAGAGCTCCAGTTTAAAGTTTGAACAAGCCAACTTTAAGCCATGGGAAGGG	229
QY	2389	GAGAGCATATCATGATGTGGGTATGAGAGATCTTGAGCTTACACTTGGAAATGGGAAGCT	2448
Db	228	GAGAGCATATCATATGTGGCTTATGGAGAGATCTTAGAGCTCAGCGTTGGAAATGGTAAAGCT	169
QY	2449	CGTGGTTGACATTAACAACACTAGGCAAAACCCAAATGSCATGTC	2490
Db	168	CATAGTTGAGATGAATATATCATGAAGAACTGTATATGTCATGTC	127

RESULT 9

LOCUS	AV551486	585 bp	mRNA	linear	EST 06-SEP-2000
DEFINITION	AV551486 Arabidopsis thaliana roots	Columbia	Arabidopsis	thaliana	
CDNA	clone R2127d10R 5'	mRNA sequence.			
ACCESSION	U5551486				

VERSION AV551486.1 GI:8722899
KEYWORDS EST.

ORGANISM
Aradidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE 1 ; eutrosids 11; Brassicales; Brassicaceae; Arabidopsis.
(bases 1 to 585)

JOURNAL
OF
BIOLOGICAL CHEMISTRY
DNA RES. 7, 175-180 (2000)
20363093

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"

```

```

//Trace="R2127d10R"
//clone="R2127d10R"
//clone_type="roots"
//clone_lib="Arabidopsis thaliana roots Columbia"

```

BASE COUNT	146	142	138	159
ORIGIN	a	c	g	t

Query Match	13.2%	Score 356	DN 9	Length 585
Best Local Similarity	100.0%	Pred. No. 9.4e-47		
Matches 356; Conservative	0	Mismatches 0	Indels 0	Gaps 0

OY 1015 CTTCATCAGCTACGTTACACATCTTTACCCACCTTATCAAGCTCTCCGGAAGT 1074
 DB 1 CTTCATCAGCTACGTTACACATCTTTACCCACCTTATCAAGCTCTCCGGAAGT 60
 OY 1075 TTCGAGGGGTTGTTGTTGTTTCCCGATGACTTCGATGAGCTACGTCGACATCAGC 1134
 DB 61 TTCGAGGGGTTGTTGTTGTTTCCCGATGACTTCGATGAGCTACGTCGACATCAGC 120
 OY 1135 AGCTCAGCTGCGCGCGCGCGAGCAGCGCGCGAGAGAAAGCCGCTGCGAGAGAAAGGA 1194
 DB 121 AGCTCAGCTGCGCGCGCGCGAGCAGCGCGCGAGAGAAAGCCGCTGCGAGAGAAAGGA 180
 OY 1195 GAATATCGAGAGAGAGAAAGCGAGAGAGCTGATGCTGCGAGAGATATACACTGGGCT 1254
 DB 181 GAATATCGAGAGAGAGAAAGCGAGAGAGCTGATGCTGCGAGAGATATACACTGGGCT 240
 OY 1255 TAGAGCTCAGAGGATTTATATTTGCTTAAACATTTGATATATATGAAGTCTTAAAGC 1314
 DB 241 TAGAGCTCAGAGGATTTATATTTGCTTAAACATTTGATATATATGAAGTCTTAAAGC 300
 OY 1315 TCTTTGTGTTGAAGCTGTTGGTGTGTTGAAGAGATGTTACTTATTCGCAAG 1370
 DB 301 TCTTTGTGTTGAAGCTGTTGGTGTGTTGAAGAGATGTTACTTATTCGCAAG 356
 RESULT 10
 T42881 514 bp mRNA linear EST 07-JAN-1998
 LOCUS 6144 Lambda-PRL2 Arabidopsis thaliana cDNA clone 119H677, mRNA
 DEFINITION
 T42881
 VERSION T42881.1 GI:2758126
 KEYWORDS
 SOURCE Arabidopsis thaliana (chale crees)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 euroids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 514)
 Newman,T., deBull,J., F.J., Green,P., Keesstra,K., Kende,H., McIntosh
 L., Ohlrogge,J., Raithel,N., Somerville,S., Thomasow,M., Retzel
 E. and Somerville,C.
 Gene galeore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 7846151
 COMMENT On Jan 7, 1998 this sequence version replaced gi:947520.
 CONTACT: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313c@msu.edu
 Seg primer: T7.
 Location/Qualifiers
 1..514
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="var Columbia"
 /db_xref="taxon:3702"
 /clone="119H677"
 /clone_lib="Lambda-PRL2"
 /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) etiolated plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 elliques. The vector is BRL's lambda Zip-Lox. The cDNA

BASE COUNT 126 a 113 c 123 g 132 t 20 others
 ORIGIN
 Query Match 13.1%; Score 352.8; DB 14; Length 514;
 Best Local Similarity 92.8%; Pred. No. 3.1e-46;
 Matches 397; Conservative 0; Mismatches 27; Indels 4; Gaps 3;
 OY 929 AGAAGAGAGATTTCTTCTTCTGATTCGAGGAGAGAAAGCGTATTCCTCGTGAGAC 988
 DB 1 AGAAGAGAGATTTCTTCTTCTGATTCGAGGAGAGAAAGCGTATTCCTCGTGAGAC 60
 OY 989 TAACCTTCACCTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1048
 DB 61 TAACCTTCACCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 120
 OY 1049 ACCATTCAGAGCTCTCTCTCGGAGGTTTGAAGGCTTGTGTTTCCCGATGACT 1108
 DB 121 ACCATTCAGAGCTCTCTCTCGGAGGTTTGAAGGCTTGTGTTTCCCGATGACT 180
 OY 1109 TCGGATGAGCTACGTCGATCAGCAGCTGACGCTGCGGCGCGCAGCAGCGCGAGG 1168
 DB 181 TCGGATGAGCTACGTCGATCAGCAGCTGACGCTGCGGCGCGCAGCAGCGCGAGG 239
 OY 1169 AGAAGCCGCTGTCGAGAGAGAGAGAGATATTCGAGAGAGAGAGAGAGAGAGAGAG 1228
 DB 240 AGAAGCCGCTGTCGAGAGAGAGAGAGATATTCGAGAGAGAGAGAGAGAGAGAGAG 299
 OY 1229 GTAGCTGCGAGAGATATACCTGCGCTTAAAGCTCAGCTATATATATTTGGCTTAAACAT 1288
 DB 300 GTAGCTGCGAGAGATATACCTGCGCTTAAAGCTCAGCTATATATATTTGGCTTAAACAT 359
 OY 1289 TGTGAT 1345
 DB 360 TGTGAT 419
 OY 1346 GAAGATGG 1353
 DB 420 GAAGATGG 427
 RESULT 11.
 BH673471 773 bp DNA linear GSS 19-FEB-2002
 LOCUS BOHXG46TR BO_2_3_KB Brassica oleracea genomic clone BOHXG46.
 DEFINITION
 BOHXG46TR BO_2_3_KB Brassica oleracea genomic clone BOHXG46.
 genomic survey sequence.
 BH673471
 VERSION BH673471.1 GI:18740935
 KEYWORDS
 SOURCE GSS.
 ORGANISM Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 euroids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 773)
 Town,C.D., Van Aken,S., Uteyback,T., Koo,H. and Frazer,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-0523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seg primer: TR
 Classes: sheared ends.
 Location/Qualifiers
 1..773
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"

AUTHORS Pih,K.T., Park,J.M., Jang,H.J., Kang,S.G., Piao,H.L. and Hwang,I.
TITLE EST of salt inducible mRNA in Arabidopsis thaliana
JOURNAL Unpublished
COMMENT Contact: Kyeong Tae Pih
 Department of Plant Molecular Biology
 Gyeongang National Univ., Plant Molecular Biology and
 Biotechnology Research Center
 Jinju, Kyungnam 660-701, Korea
 Tel: 82-591-751-5193
 Fax: 82-591-759-9363
 Email: ihwang@nongae.gsnu.ac.kr.

FEATURES
source
 1. 400
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="OS158"
 /dev_stage="7 day-old"
 /clone_lib="NaCl-treated Arabidopsis subtraction library"
 /note="Subtracted cDNA library from salt (NaCl)-treated whole plants"

BASE COUNT 96 a 124 c 57 g 121 t 2 others
ORIGIN

Query Match 11.3%; Score 303.8; DB 14; Length 400;
 Best Local Similarity 98.8%; Pred. No. 2e-38;
 Matches 327; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

1792 CTTTATCAGCGCTTTCAAGTCCCATCCATCTTACCAAGTTAGCCGCTTCTTCATCA 1851
 1 CTTTATCAGCGCTTTCAAGTCCCATCCATCTTACCAAGTTAGCCGCTTCTTCATCA 60
 1852 TTCGCGAGTCCTTCGCGGTGAACCAATTAACAAGATGCTCTCAATTCCTTC 1911
 61 TTCGCGAGTCCTTCGCGGTGAACCAATTAACAAGATGCTCTCAATTCCTTC 120
 1912 CTCAGAAATGATGGCAATTCCTTCCTTCCTCCCTCAGAAATCTCAAGTGTCCA 1971
 121 CTCAGAAATGATGGCAATTCCTTCCTTCCTCCCTCAGAAATCTCAAGTGTCCA 180
 1972 GTTACCCCAACCGGTCTCATGCGGATCTTAAGAACCCGAACCGTGTCTAATGCGAA 2031
 181 GTTACCCCAACCGGTCTCATGCGGATCTTAAGAACCCGAACCGTGTCTAATGCGAA 240
 2032 TCTATGCTAAGCAATCCATCGGCATTTGCTAAACATCAATGAGCGTC-TTTAATATATC 2089
 241 TCTATGCTAAGCAATCCATCGGCATTTGCTAAACATCAATGAGCGTC-TTTAATATATC 300
 2090 CTTTCTATGCGGTTTCTGCACTGCTAGTCC 2120
 301 CTTTCTATGCGGTTTCTGCACTGCTAGTATTC 331

RESULT 14 307 bp mRNA linear EST 23-APR-1997
AA389830
LOCUS OS198 NaCl-treated Arabidopsis subtraction library Arabidopsis thaliana cDNA 5', mRNA sequence.
DEFINITION AA389830 GI:2042816
ACCESSION AA389830
VERSION AA389830.1
KEYWORDS EST.
SOURCE Arabidopsis thaliana (chale crese)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
AUTHORS Pih,K.T., Park,J.M., Jang,H.J., Kang,S.G., Piao,H.L. and Hwang,I.
TITLE EST of salt inducible mRNA in Arabidopsis thaliana
JOURNAL Unpublished
COMMENT Contact: Kyeong Tae Pih
 Department of Plant Molecular Biology

Gyeongang National Univ., Plant Molecular Biology and
 Biotechnology Research Center
 Jinju, Kyungnam 660-701, Korea
 Tel: 82-591-751-5193
 Fax: 82-591-759-9363
 Email: ihwang@nongae.gsnu.ac.kr
 Seq primer: T3
 High quality sequence stop: 307.

FEATURES
source
 1. 307
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /dev_stage="7 day-old"
 /clone_lib="NaCl-treated Arabidopsis subtraction library"
 /note="Subtracted cDNA library from salt (NaCl)-treated whole plants"

BASE COUNT 71 a 197 c 41 g 96 t 2 others
ORIGIN

Query Match 10.7%; Score 288.8; DB 9; Length 307;
 Best Local Similarity 97.0%; Pred. No. 5e-36;
 Matches 293; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1792 CTTTATCAGCGCTTTCAAGTCCCATCCATCTTACCAAGTTAGCCGCTTCTTCATCA 1851
 1 CTTTATCAGCGCTTTCAAGTCCCATCCATCTTACCAAGTTAGCCGCTTCTTCATCA 60
 1852 TTCGCGAGTCCTTCGCGGTGAACCAATTAACAAGATGCTCTCAATTCCTTC 1911
 61 TTCGCGAGTCCTTCGCGGTGAACCAATTAACAAGATGCTCTCAATTCCTTC 120
 1912 CTCAGAAATGATGGCAATTCCTTCCTTCCTCCCTCAGAAATCTCAAGTGTCCA 1971
 121 CTCAGAAATGATGGCAATTCCTTCCTTCCTCCCTCAGAAATCTCAAGTGTCCA 180
 1972 GTTACCCCAACCGGTCTCATGCGGATCTTAAGAACCCGAACCGTGTCTAATGCGAA 2031
 181 GTTACCCCAACCGGTCTCATGCGGATCTTAAGAACCCGAACCGTGTCTAATGCGAA 240
 2032 TCTATGCTAAGCAATCCATCGGCATTTGCTAAACATCAATGAGCGTC-TTTAATATATC 2091
 241 TCTATGCTAAGCAATCCATCGGCATTTGCTAAACATCAATGAGCGTC-TTTAATATATC 300
 2092 TT 2093
 301 TT 302

RESULT 15 602 bp DNA linear GSS 13-DEC-2002
B2476824/c
LOCUS BONBH01TR B0.1 6.2 KB tot Brassica oleracea genomic clone BONBH01,
DEFINITION genomic survey sequence.
ACCESSION B2476824
VERSION B2476824.1
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
AUTHORS Town,C.D., Van Aken,S., Utecherack,T., Koo,H. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished
COMMENT Other GSSs: BONBH01TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-8523
 Fax: 301-838-0208

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 19:12:39 ; Search time 153 Seconds
(without alignments)
7751.611 Million cell updates/sec

Title: US-09-995-938A-2

Perfect score: 2687

Sequence: 1 actattgttcgcaattc.....attattgtgataaaaaa 2687

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 563978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing filter 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/pdata/1/ina/5A_COMB.seq:*
2: /cgn2_6/pdata/1/ina/5B_COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A_COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B_COMB.seq:*
5: /cgn2_6/pdata/1/ina/PCUS_COMB.seq:*
6: /cgn2_6/pdata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66.6	2.5	7218	1	US-08-232-463-14 Sequence 14, Appl
2	50.4	1.9	12730	2	US-09-004-838-91 Sequence 91, Appl
3	48.4	1.8	19124	2	US-08-487-8268-13 Sequence 13, Appl
4	47.8	1.8	15062	2	US-09-004-838-89 Sequence 89, Appl
5	46.6	1.7	2435	3	US-09-306-593-1 Sequence 1, Appl
6	46.2	1.7	7218	1	US-08-332-463-14 Sequence 14, Appl
7	46.2	1.7	9510	4	US-07-991-8678-8 Sequence 256, App
8	45.2	1.7	1511	1	US-08-107-7558-8 Sequence 8, Appl
9	45.2	1.7	1511	1	US-08-544-332-8 Sequence 8, Appl
10	45.2	1.7	1511	2	US-09-370-861A-8 Sequence 8, Appl
11	45.2	1.7	1511	4	US-09-426-290-1 Sequence 124, Appl
12	44.6	1.7	168753	4	US-09-004-838-124 Sequence 124, Appl
13	44.4	1.7	12793	4	US-08-998-416-1137 Sequence 1137, Ap
14	43.4	1.6	636	3	US-08-998-416-1137 Sequence 1, Appl
15	43.4	1.6	14507	3	US-09-660-299-1 Sequence 1, Appl
16	43.4	1.6	14507	4	US-09-435-377-1 Sequence 1, Appl
17	43.4	1.6	14507	3	US-08-998-416-1137 Sequence 186, App
18	43.2	1.6	615	4	US-09-521-511C-10 Sequence 10, Appl
19	43.2	1.6	4262	4	US-09-041-886-10 Sequence 10, Appl
20	42.8	1.6	3715	4	US-08-916-421B-1 Sequence 1, Appl
21	42.8	1.6	1664976	4	US-09-601-198-3 Sequence 3, Appl
22	42.6	1.6	1308	4	US-09-801-861-3 Sequence 3, Appl
23	42.2	1.6	53332	4	US-08-998-416-534 Sequence 534, App
24	42	1.6	854	3	US-07-867-106-2 Sequence 2, Appl
25	41.8	1.5	5852	1	US-08-998-416-595 Sequence 287, App
26	41.6	1.5	860	3	US-09-601-198-60 Sequence 60, Appl
27	41.4	1.5	15016	4	US-09-601-198-60 Sequence 60, Appl

28	41.4	1.5	168575	4	US-09-426-290-1 Sequence 1, Appl
29	41.2	1.5	340	1	US-08-628-417-6 Sequence 6, Appl
30	41.2	1.5	732	3	US-08-998-416-1036 Sequence 1036, Ap
31	41.2	1.5	1798	4	US-09-797-906-1 Sequence 1, Appl
32	41.2	1.5	19124	2	US-08-487-8268-13 Sequence 13, Appl
33	41	1.5	4185	4	US-09-417-485D-7 Sequence 7, Appl
34	41	1.5	5852	1	US-07-867-106-2 Sequence 2, Appl
35	40.8	1.5	658	3	US-08-998-416-595 Sequence 595, App
36	40.8	1.5	1966	4	US-09-835-811-1 Sequence 1, Appl
37	40.6	1.5	4673	1	US-07-638-431-1 Sequence 1, Appl
38	40.6	1.5	4673	5	PCT-US92-00018-1 Sequence 1, Appl
39	40.6	1.5	6243	2	US-09-056-075-1 Sequence 1, Appl
40	40.4	1.5	7832	4	US-09-004-838-94 Sequence 94, Appl
41	40.4	1.5	10007	4	US-09-410-464-13 Sequence 13, Appl
42	40.4	1.5	10640	4	US-09-417-485D-5 Sequence 5, Appl
43	40.2	1.5	658	3	US-08-998-416-595 Sequence 595, App
44	40.2	1.5	872	3	US-08-998-416-487 Sequence 487, App
45	40.2	1.5	6124	4	US-08-213-419B-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
CLASSIFICATION: 415
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 INMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 846-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZpt-Fls
US-08-232-463-14
Query Match 2.5%; Score 66.6; DB 1; Length 7218;

Best Local Similarity 3.8%; Pred. No. 2.1e-07;
Matches 15; Conservative 233; Mismatches 147; Indels 0; Gaps 0;

QY 1793 TTTCATCAGCCTTCAAGTCCCATCCATCCATCAAGTACGCCCTCTTCATCAT 1852
1067 YY 1126
QY 1853 TCCCGAGTCTTCTCGCGGGAACCAATACAAATGCTCTTCACTTCTCCCTTCC 1912
1127 YY 1186
QY 1913 TCAGAAATGGTGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1972
1187 YY 1246
QY 1973 TTACCCACCGGCTCTCATCCGCACTTCAAGAACCGGAACCGTCTCACTGGGAT 2032
1247 YY 1306
QY 2033 CTATCGCTAAGCAATCCATGCCATTCCTTAAACATCAATGCGCTTTTAATTCCTT 2092
1307 YY 1366
QY 2093 TCTATGGGCTTCTCGACCTGCTAGTCCGACATCGCACAGTTTCATACCTGGCTA 2152
1367 YY 1426
QY 2153 CTATACCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2187
1427 YYYYYYYYGTACCAAAATCTTCTATCTCTTAACT 1461
Db

RESULT 2
US-09-004-838-91/c
Sequence 91, Application US/09004838
Patent No. 6350933
GENERAL INFORMATION:
APPLICANT: Michelmore, Richard W.
APPLICANT: Shen, Kathy
APPLICANT: Meyers, Blake
TITLE OF INVENTION: Procedures and Materials for
TITLE OF INVENTION: Confering Pest Resistance in Plants
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 09-JAN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/781,734
FILING DATE: 10-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Elmhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 023070-078810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 12730 base pairs
TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: 1.12730

LOCATION: 1.12730

OTHER INFORMATION: /note= "RG2C"

US-09-004-838-91

Query Match

Best Local Similarity 47.9%; Pred. No. 0.0059;

Matches 205; Conservative 0; Mismatches 221; Indels 2; Gaps 2;

QY 173 GCCGAACCTTGTTATGGAGCAAACTGTAACCCCTTTTCTTTTATGTTCAATGA 232
7038 GCGAAACGTAAGATATAGAAATATATATATCTTATGACCTTTGGCTTAATATA 6979
QY 233 ATACAAAGTTTGTTATGATATACATTAATATGATGACCGAGCATTAATCCAAATTT 292
6978 TATATATGTAAGTTAGAGAAAGAAAGTCAAGTAAGCAAGAAATTTGGCAAGTT 6919
QY 293 GATATATAGATCTAAGCTTAAATCAACATGTAACCAACTAAATCTTATAGAAC 351
6918 TGAATATATAGAACTTACTTCTTCTTAAAGACATTAAGAAATTAATATATATACTA 6859
QY 352 ATAGTAAGGATATTCACCAATCTTATATCATTTTGAAGTACGAAGAAAGTAAATA 411
6858 TTCTTAAAGACCTTGTGAATTAATATATATATATATATATATATATATATATATAT 6799
QY 412 AGAGAGAGCGATGATCATCACTAATCAAGCAAAAGTAAAGTAAATGTTCTTAAG 471
6798 TAAAGAACTCAATTTGAAGCAAACTGAAATTTTAAATAA-TTAAATTCATGATTTAA 6740
QY 472 TGAGATTGTATGCAAGAAAGAAAGTAAATTTTAAATGATATATATATATATATAT 531
6739 ATTAATTTTCAATTTCAATTAATTTGACATAATATATATATATATATATATATATACAA 6680
QY 532 ATACACGGCTTCACTGATTAATAATTAATTTTAAACCAATCAATGATGTTTTC 591
6679 GCAATATTAATTAATTAAGATTAACATTAATATTTTATTAAGTTGTAATTTTATATA 6620
QY 592 GGTACAC 599
6619 TATATATAC 6612
Db

RESULT 3

US-08-487-8268-13

Sequence 13, Application US/084878268

Patent No. 593827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chean

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhuang

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knodde Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NH121,001CPL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 1.8%; Score 48.4; DB 2; Length 19124;
Best Local Similarity 50.6%; Pred. No. 0.025;
Matches 170; Conservative 0; Mismatches 161; Indels 5; Gaps 2;

QY 246 TTATGATACATTAATATGATGACCCGCAATTATCCAAATTGGATTTGATATC 305
DB 15614 TTTTAT 15673
QY 306 TAAAGCTTAAATCAACATGTAACCACTTAATATCTTATGACATGTAATGCTAT 365
DB 15614 ATTAAAT 15732
QY 366 TCACCAATCTTAT 15732
DB 15723 --ATAAAT 15790
QY 426 TACATCACTATATGACGACAAAGATGCAAAATGTTCTTAAAGTGATTTGATATC 485
DB 15791 TACATCACTATATGACGACAAAGATGCAAAATGTTCTTAAAGTGATTTGATAT 15850
QY 486 AAGAAAGAGAT 543
DB 15851 AAAT 15910
QY 544 CACTGAT 579
DB 15911 AATGAAAT 15946

RESULT 4
US-09-004-838-89/c
Sequence 89, Application US/09004838
Patent No. 6350933
GENERAL INFORMATION:
APPLICANT: Michelmore, Richard W.
APPLICANT: Shen, Kathy
APPLICANT: Meyers, Blake
TITLE OF INVENTION: Procedures and Materials for
TITLE OF INVENTION: Confering Pest Resistance in Plants
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESSES:
ADDRESSES: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,838
FILING DATE: 09-JAN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/781,734
FILING DATE: 10-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Einhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 023070-078810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 15062 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: 1.15062
LOCATION: 1.15062
OTHER INFORMATION: /note="RG2B"
US-09-004-838-89

Query Match 1.8%; Score 47.8; DB 4; Length 15062;
Best Local Similarity 47.4%; Pred. No. 0.032;
Matches 203; Conservative 0; Mismatches 223; Indels 2; Gaps 2;

QY 173 GCCGAACCTGTATATGAGCAAACTGTAACCCCTTTTCTTATATGTTCAATGACT 232
DB 9453 GGCAGAACCTGTAAT 9394
QY 233 ATACAGTTTGTATGAT 292
DB 9393 TAT 9335
QY 293 GGAT 352
DB 9334 GTTAT 9275
QY 353 TAT 411
DB 9274 ATGCAAT 9215
QY 412 AAGAGAGCGCATGTAT 471
DB 9214 TTAAGAACTCAATTTGAAGCAAACTGAATAATTTTAAATATATATATATATAT 9155
QY 472 TGAGATTTGTATGACAAAGAAAGATATATATATATATATATATATATATATATAT 531
DB 9154 ATTAATTTTAT 9095
QY 532 ATCAGCGGCTTCACTAT 591
DB 9094 GCAAT 9035
QY 592 GGTACAC 599
DB 9034 TATATATAC 9027

RESULT 5
US-09-306-593-1
Sequence 1, Application US/09306593
Patent No. 6184018
GENERAL INFORMATION:
APPLICANT: Li, Xin-Liang
APPLICANT: Ljungdahl, Lars G.
APPLICANT: Chen, Huizhong
APPLICANT: Ximenes, Eduardo A.

TITLE OF INVENTION: Beta-glucosidase Coding Sequences and Protein from
FILE REFERENCE: 31-98us
CURRENT APPLICATION NUMBER: US/09/306,593
CURRENT FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: US 60/084,494
EARLIER FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2435
TYPE: DNA
ORGANISM: Orpinomyces sp. PC-2
FEATURE:
NAME/KEY: CDS
LOCATION: (39)..(2009)
FEATURE:
NAME/KEY: mat peptide
LOCATION: (87)..(2009)
US-09-306-593-1

Query Match 1.7%; Score 46.6; DB 3; Length 2435;
Best Local Similarity 45.7%; Pred. No. 0.024;
Matches 163; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 222 TTCAATGACTATACAGCTTTGGTATGAATACATAAATATGATGACCCAGCAATTA 281
DB 2079 TTTATTCCTCTTTGCTATAGAGTAATAATTAATTTATTAATTAATTT 2138
QY 282 ATCAAAATTTGATATAGATACCTAAAGCTTAAATCAACATGTAACCAACTAAATC 341
DB 2139 ATAGAAATTAATCTTTTGAATCATTAATTAATTAATTAATTAATTAATTAATGA 2198
QY 342 TTTATAGAACATGATTAATGATATACCAATCTTATATCATTTGTAAGTAGAGAA 401
DB 2199 ATATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2258
QY 402 GGTAAAAAAGAGAGACCGAGTACATACCACTAATCAGACAAAAGTCAAAAT 461
DB 2259 AAGCAAGAAACCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2318
QY 462 GTTCTAAGTGAATTTGTATGCAAGAAAAGTAAATTTTAATGAATATATCAT 521
DB 2319 AATAAGCTTGTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2378
QY 522 TATGATGTTAATCAGACCGCTTCACTGATTAATTAATTAATTAATTAATTAAT 578
DB 2379 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2435

RESULT 6

US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHREIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 856-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-Fls
US-08-232-463-14

Query Match 1.7%; Score 46.2; DB 1; Length 7218;
Best Local Similarity 3.4%; Pred. No. 0.056;
Matches 9; Conservative 157; Mismatches 95; Indels 0; Gaps 0;

QY 1111 GGATGAGCTACGTGACATCAGCAGCTGCAGCTGGCGGCGGAGAGCGGAGAG 1170
DB 1317 RRR 1258
QY 1171 GAACCGCTGCGAGAGAAAGGGAATATCGAGAGAGAGAGAGAGAGAGAGAGCT 1230
DB 1257 RRR 1198
QY 1231 AGTCGAGATATACACTGGGCTTAGAGCTCAAGTGATATTAATTTGCCATAACATTG 1290
DB 1197 RRR 1138
QY 1291 TGATTAATGAAGTCTTAAGCTCTTTGTGTTGAAGCTGTTGGTTGAAGAAGA 1350
DB 1137 RRR 1078
QY 1351 TGGTACTACTATGCAAGCT 1371
DB 1077 RRRRRRRRRRRATCGCAAGCT 1057

RESULT 7

US-09-453-702B-256
Sequence 256, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blatner, Frederick R.
Burland, Valerie
Peral, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-Dec-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296, 95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 256:
SEQUENCE CHARACTERISTICS:
LENGTH: 9510 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 256:
US-09-453-702B-256

Query Match 1.7%; Score 46.2; DB 4; Length 9510;
Best Local Similarity 51.2%; Pred. No. 0.066;
Matches 108; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 276 CAATTAATCAAAATTTGGATTGATTAAGTAAAGTTAAATCAACATGTATACCAACT 335
DB 1185 CAACATATTTACATATTTGATATCAAAATGAAATATCAAAATGATATGATATCCAGACT 1244
QY 336 AAATCTTATAGACATAGTAATGATATCAACCAATCTTATATCATTTGTAAAGTAC 395
DB 1245 TGACTTTTAACTTAACGTTTAAATCTTACACAGAGTTCAAAATGATATTAATCTAC 1304
QY 396 GAAAGAGTAAAG 455
DB 1305 AGTAG 1364
QY 456 AAAATTTGTTCTAAAGTGAATTTGATGCA 486
DB 1365 GGATGATTTCAACTGAGAGATTTAACA 1395

RESULT 8
US-07-991-867B-8
Sequence 8, Application US/07991867B
Patent No. 5476781
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
FILING DATE: 12-DEC-1992
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UP14.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amaseta moorei entomopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-07-991-867B-8

Query Match 1.7%; Score 45.2; DB 1; Length 1511;
Best Local Similarity 44.7%; Pred. No. 0.043;
Matches 176; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 181 TTGTTATGGAGAACTGCTAAACCCCTTTTCCTTTATGTTCAATGACTATACAGT 240
DB 737 TTGTTTGGCAAGAAACATAGAGCAATATTAAATCTATGACATTTTATTTATTTAT 796
QY 241 TTGGTTATGATATCATTAATATATGATGACCCACCAATTAATCCAAATTTGATATTA 300
DB 797 TGATATATTTTTCACAAAATTAATATCATGMAAAAAATTAATATCAAAATGCA 856
QY 301 GATACCTAAGCTTAATCAACATGTAACCAACCTAATCTTATAGAACATAGTAAAT 360
DB 857 TTTACTAAATTTGATATATTTTAAATATTAATTTTAAATATTTATTAATTTAAAAAAT 916
QY 361 GGATTCACCAATCTTATATCATTTGTAAGTACGAGAGAGTAAAGAGAGAGAGAGAGAG 420
DB 917 AATTAATTAACAGAGAAATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 976
QY 421 CAGGTACATACACTATATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 977 AGAATTCATTA 1036
QY 481 TATGCAAGAAAAAGTGAATTTTAAATGAATATATATATATATATATATATATATAT 540
DB 1037 AAGTTATATATTTCAAAATTTAAATTTATTAATTTTAAATTTTATTAACAAAATTA 1096
QY 541 CTCACCTGTAATAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 574
DB 1097 TATTAATATTTGATATATATCTTATTAACAAAATTT 1130

RESULT 9
US-08-107-755A-8
Sequence 8, Application US/08107755A
Patent No. 5721352

GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5721352e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSER: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UFI14.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Ambacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-08-107-755A-8

Query Match 1.7%; Score 45.2; DB 1; Length 1511;
Best Local Similarity 44.7%; Pred. No. 0.043; Indels 0; Gaps 0;
Matches 176; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

Qy 181 TTGTTATGAGCAAACTCGTAAACCCCTTTTCCTTTTATGTTCAATGAACTATACACT 240
Db 737 TTGTTTTCGCAAAAACATAGAGCAATATTATTCATTCGACATTTTITTTTATATT 796
Qy 241 TTTGGTTATGAACTAAATAATGATGAGCCAGCAATTAATCCAAAATTGGATATTA 300
Db 797 TGATATATTTTTCAAAATAATTAATCAATGAAAAAATAAATAATATCAAAATGGA 856
Qy 301 GATTAAGCTTAATAATCAATGAACTAAATTAATTAATGAACTAGTAAT 360
Db 857 TTTACTAAATTTCTGATATATTTTAAATATTTTAAATATTAATTTTAAAAAAAT 916
Qy 361 GGTATTCACCAATCTTATATCATTTGTAGTACGAGAGGTAATAAAGAGAGAC 420

Db 917 AATAATAACAGAGATATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 976
Qy 421 CAGTGTACATACACTAATACAGCAAAAGTACTCAAAATGTTCTAAAGTGAGATTG 480
Db 977 AGAATGCAATTA 1036
Qy 481 TATGCAAGAAAAAGGATATATTTTAAATGATATATATATATATATATATATATAT 540
Db 1037 AGTTTATATATTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1096
Qy 541 CTTCAGCTATATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 574
Db 1097 TATTAATATTTAGATATATCTTATTAACAAAAT 1130

RESULT 10
US-08-544-332-8
Sequence 8, Application US/08544332
Patent No. 5935777
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSER: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UFI14.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Ambacta moorei entemopoxvirus
FEATURE:

NAME/KEY: CDS
LOCATION: Complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: Complement (234..782)
NAME/KEY: CDS
LOCATION: 852..1511
US-08-544-332-8

Query Match
Best Local Similarity 44.7%; Score 45.2; DB 2; Length 1511;
Matches 176; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

1.7%; Score 45.2; DB 2; Length 1511;
Best Local Similarity 44.7%; Pred. No. 0.043;
Matches 176; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

181 TTGTTATGGGACAACTCGTAAACCCCTTTTCCTTTATGTTCAATGACTATCAACT 240
737 TTGTTGGCCGAGAAACATAGACCAATTAATTCATGCACTTTTATTTATTTAT 796
241 TTGTTATGAAATACATTAATTAATGATGACCCAGCAATTAATCCAAATTTGATATTA 300
797 TGATATATTTTTCATAAAATTAATCAATGAAAAAATTAATTAATCAAAATGGA 856
301 GATCTAAAGCTTAAATCAACATGTAACCAATTAATCTTTATAGAACATAGTAAT 360
857 TTCTACTAAATTCGATATTAATTTTAATTAATTTTAATTAATTTAAATTTAA 916
361 GGTATTCACCAATCTTTATCATTTGTAAGTACGAGAGAGAGAGAGAGAGAG 420
917 AATATATAACAGAGATTAATTTATTAATTTAAATTAATTAATTAATTTAA 976
421 CAGGTACATACAACTAAATCAGACAAAGTAGTCAAAATTTCTTAAGTGAATTTG 480
977 AGAATTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1036
481 TATGCAAGAAAAAGTGAATTTTAAATGAATATATATGATTAATGATTAATCAACGG 540
1037 AAGTTATATATTTCAAAATTTAAATTTATTAATTTAAATTTTATTAACAAATTTAA 1096
541 CTTCACGTATTAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 574
1097 TATAACATATTTAGATATATCTTATTAACAAAAAT 1130

RESULT 11
US-09-370-861A-8
Sequence 8, Application US/09370861A
Patent No. 6410221
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 6410221el Entomopoxvirus Expression System
CURRENT APPLICATION NUMBER: US/09/370, 861A
CURRENT FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: US 07/991,867
PRIOR FILING DATE: 1992-12-07
PRIOR APPLICATION NUMBER: US 08/107,755
PRIOR FILING DATE: 1993-08-19
PRIOR APPLICATION NUMBER: WO 92/14818
PRIOR FILING DATE: 1992-02-12
PRIOR APPLICATION NUMBER: US 07/827,685
PRIOR FILING DATE: 1992-01-30
PRIOR APPLICATION NUMBER: US 07/657,584
PRIOR FILING DATE: 1991-02-19
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin version 3.1
SEQ ID NO 8
LENGTH: 1511
TYPE: DNA
ORGANISM: Ambacta moorei entomopoxvirus
US-09-370-861A-8

Query Match
Best Local Similarity 44.7%; Score 45.2; DB 4; Length 1511;
Matches 176; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

1.7%; Score 45.2; DB 4; Length 1511;
Best Local Similarity 44.7%; Pred. No. 0.043;
Matches 176; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

181 TTGTTATGGGACAACTCGTAAACCCCTTTTCCTTTATGTTCAATGACTATCAACT 240
737 TTGTTGGCCGAGAAACATAGACCAATTAATTCATGCACTTTTATTTATTTAT 796
241 TTGTTATGAAATACATTAATTAATGATGACCCAGCAATTAATCCAAATTTGATATTA 300
797 TGATATATTTTTCATAAAATTAATCAATGAAAAAATTAATTAATTAATTAATGGA 856
301 GATCTAAAGCTTAAATCAACATGTAACCAATTAATCTTTATAGAACATAGTAAT 360
857 TTCTACTAAATTCGATATTAATTTTAATTAATTTTAATTAATTTAAATTTAA 916
361 GGTATTCACCAATCTTTATCATTTGTAAGTACGAGAGAGAGAGAGAGAGAG 420
917 AATATATAACAGAGATTAATTTATTAATTTAAATTTAAATTTAAATTTTAA 976
421 CAGGTACATACAACTAAATCAGACAAAGTAGTCAAAATTTCTTAAGTGAATTTG 480
977 AGAATTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1036
481 TATGCAAGAAAAAGTGAATTTTAAATGAATATATATGATTAATGATTAATCAACGG 540
1037 AAGTTATATATTTCAAAATTTAAATTTATTAATTTTAAATTTTATTAACAAATTTAA 1096
541 CTTCACGTATTAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 574
1097 TATAACATATTTAGATATATCTTATTAACAAAAAT 1130

RESULT 12
US-09-426-290-1/c
Sequence 1, Application US/09426290
Patent No. 6410712
GENERAL INFORMATION:
APPLICANT: Berglund Rafi Oiafedoctir
APPLICANT: Jeffrey Guilcher
TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
FILE REFERENCE: 2345.2001-000
CURRENT APPLICATION NUMBER: US/09/426,290
CURRENT FILING DATE: 1999-10-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FaeltSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 168575
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (21181)...(21403)
NAME/KEY: CDS
LOCATION: (95252)...(95430)
NAME/KEY: CDS
LOCATION: (101753)...(101996)
NAME/KEY: CDS
LOCATION: (110324)...(110439)
NAME/KEY: CDS
LOCATION: (124058)...(124278)
NAME/KEY: CDS
LOCATION: (127009)...(127130)
NAME/KEY: CDS
LOCATION: (128910)...(129139)
US-09-426-290-1

Query Match
Best Local Similarity 46.8%; Score 44.6; DB 4; Length 168575;
Matches 140; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

1.7%; Score 44.6; DB 4; Length 168575;
Best Local Similarity 46.8%; Pred. No. 0.09;
Matches 140; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

120 CAAATTAACGTCATATCAAAATTTTATGATCATCATATATATGCCCCGCAAC 179

TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1137:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1692RP
US-08-998-416-1137

Query Match 1.6%; Score 43.4; DB 3; Length 636;
Best Local Similarity 47.0%; Pred. No. 0.079;
Matches 166; Conservative 0; Mismatches 186; Indels 1; Gaps 1;

QY 226 ATGACCTATACAGTTGGTATGATCATTAATTAATGATGAGCCAGCACTTAATCC 285
DB 571 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 512
QY 286 AAAATTTGATATTAGATCTAAAGCTTAATCAACATGTAACCAACTAAATCTTTA 345
DB 511 AATAATATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATCTT 452
QY 346 TAGAACATAGTAATGATGATTCACCAATCTTTATATCATTTGTAAGTACGAAGGTA 405
DB 451 ATAAAAAGATTAATTAATTAATTAATCAAA-CATTAATATTATTAATTAATTAATTA 393
QY 406 AAAAAAGAGAGCCAGTACATCACTATCAAGCAAAAGTACGAAAATTTGTT 465
DB 392 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 333
QY 466 CTAAAGTGAATTTGATGCAAGAAAAAGTGAATTTTAATGAAATATCATTTATG 525
DB 332 TAATAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATAA 273
QY 526 ATGTTAATCAGACGGCTTCACTGTATTAATTAATTAATTAATTAATTAATTAATTA 578
DB 272 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 220

RESULT 15
US-08-785-150-1/c

Sequence 1, Application US/08785150

Patent No. 6027915

GENERAL INFORMATION:

APPLICANT: Moritz, Arvia E.

APPLICANT: Lee, Chi-Chang

APPLICANT: Thomas, James N.

TITLE OF INVENTION: Expression Augmenting Sequence Elements

Patent No. 6027915

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: Apple Macintosh

SOFTWARE: Microsoft Word for Macintosh, Version 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/785,150

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/586,509

FILING DATE: 11-JAN-96

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2841
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14507 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: No. 6027915 Relevant
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Chinese hamster
IMMEDIATE SOURCE:
CLONE: 2A5-3 lambda CHO sequence
US-08-785-150-1

Query Match 1.6%; Score 43.4; DB 3; Length 14507;
Best Local Similarity 48.9%; Pred. No. 0.47;
Matches 116; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 469 AAGTGATTTGATGCAAGAAAAAGTGAATTTTAATGAAATATATCATTTATGATG 528
DB 13710 AAAAGGATTAAGATGAATTAAGAAATTTGACCAACCAATGACATTAATTAATTA 13651
QY 529 TTAATCAGACGGCTTCACTGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 588
DB 13650 TTAAGCAGACAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13591
QY 589 TTGGTCAACAAGTAAGGAGCCACACATGAAGAACGGTCCACTGTCTCTCTCTTT 648
DB 13590 TTCCGTAAGAAAAATTAAGAGCAAAATTAATTAATTAATTAATTAATTAATTAATTA 13531
QY 649 CTTTCTCTGTATTAATTTGGTCAATCTATTTTTCATCTACACAGAAAAAATAA 705
DB 13530 CTTCAAGATGCTATTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 13474

Search completed: December 19, 2003, 20:59:43
Job time: 155 secs